

Beiträge zur Verarbeitung zweidimensionaler digitaler Signale

von
Reinhard Bernstein

Erlangen, 1981

1981 # 198

Beiträge zur Verarbeitung zweidimensionaler digitaler Signale

Dem Fachbereich Ingenieurwissenschaften
der Universität Erlangen - Nürnberg

zur Erlangung des Grades

Doktor - Ingenieur

vorgelegt von
Reinhard Bernstein

Erlangen, 1981



Als Dissertation genehmigt von der Technischen Fakultät
der Universität Erlangen - Nürnberg

Tag der Einreichung: 05.11.1980

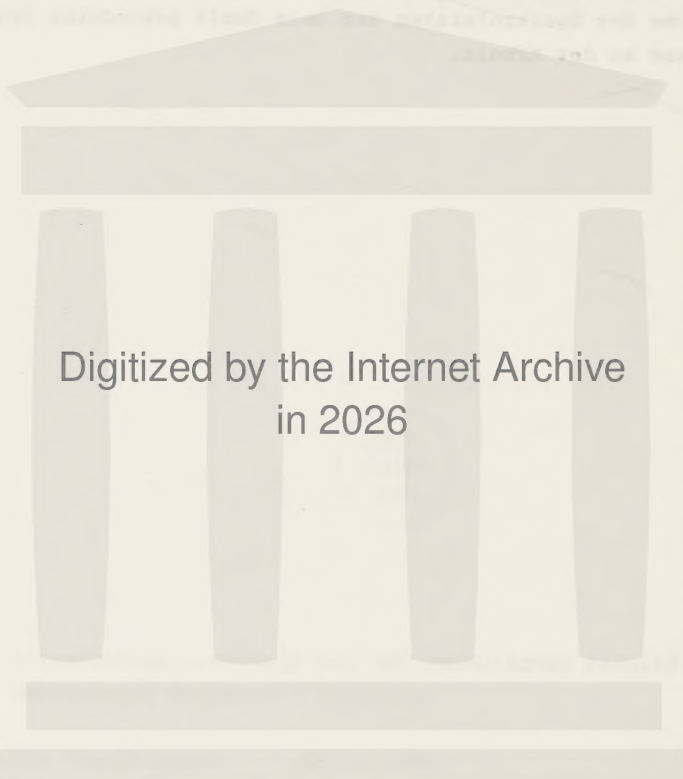
Tag der Promotion: 07.01.1981

Dekan: Prof.Dr.-Ing. H. Niemann

Berichterstatter: Prof.Dr.-Ing. W. Schüssler
Prof.Dr.rer.nat. A. Lohmann

Mein besonderer Dank gilt Herrn Prof.Dr.-Ing. W.Schüsler für die interessante Aufgabenstellung, die zahlreichen Anregungen und Hinweise und seine wertvolle Hilfe bei der Abfassung der Niederschrift.

Herrn Prof.Dr.rer.nat. A.Lohmann danke ich für die Übernahme des Zweitreferates und sein damit bekundetes Interesse an der Arbeit.



Digitized by the Internet Archive
in 2026

INHALTSVERZEICHNIS

1. Einführung und Übersicht	1
2. Lineare digitale Systeme zur 2-D Signal- verarbeitung	4
2.1 Zweidimensionale diskrete Signale	4
2.2 Definition und Eigenschaften zweidimensionaler Systeme	6
2.3 Beschreibung zweidimensionaler linearer Systeme	8
2.3.1 Beschreibung durch die Impulsantwort	9
2.3.2 Beschreibung durch die Übertragungs- funktion	11
2.3.3 Beschreibung durch die Differenzen- gleichung	12
3. Zweidimensionale nichtrekursive Filter	15
3.1 Beschreibung durch Differenzengleichung und Übertragungsfunktion	15
3.2 Spezielle Probleme	17
3.2.1 Vergrößerung des Ausgangsbildes, Randeffekte	17
3.2.2 Ortsvariante Filterung zur Vermeidung des Randeffektes	20
3.3 Implementierung eines zweidimensionalen FIR-Filters als eindimensionales System	27
3.4 Realisierung von 2-D FIR-Filtern durch schnelle Faltung	31
3.5 Zweidimensionale FFT, Implementierung an einem Minicomputer	34
3.5.1 Maschinenabhängige Größen zur Beur- teilung der Leistungsfähigkeit von Algorithmen	38
3.5.2 Zerlegung der 2-D DFT in 1-D DFT's	39
3.5.3 Gesamtrechenaufwand für die 2-D-FFT	43

3.6 Aufwandsvergleich für die Realisierung nichtrekursiver Filter an einem Minicomputer	45
4. Zweidimensionale rekursive digitale Filter	50
4.1 Grundlegende Definitionen und Eigenschaften	50
4.1.1 Differenzengleichung, Übertragungs- funktion	50
4.1.2 Verarbeitungsreihenfolge, Rekursionsrichtung	53
4.1.3 Kausalität, Semikausalität	56
4.1.4 Rekursierbarkeit, Anfangsbedingungen	59
4.2 Stabilität und Stabilitätstests	63
4.2.1 Sätze zur Stabilität	64
4.2.2 Stabilitätstests	67
4.2.2.1 Koeffizientenbedingungen für den 2-D Block 1.Grades	68
4.2.2.2 Einige bekannte numerische Testverfahren	70
4.2.2.3 Ein neuer Stabilitätstest	75
4.2.2.4 Beispiele	77
4.3 Realisierung zweidimensionaler rekursiver Filter	82
4.3.1 Realisierung von nichtkausalen und linearphasigen rekursiven Systemen	83
4.3.2 Realisierung eines Halbebenenfilters durch ein Viertelebenenfilter	87
4.3.3 Zweidimensionale Realisierung	91
4.3.4 Eindimensionale Realisierung	92
4.3.5 Realisierungsaufwand für die Durch- führung der 2-D rekursiven Filterung	99
4.4 Vergleich des Realisierungsaufwands von zweidimensionaler nichtrekursiver und rekur- siver Filterung	101
5. Entwurf zweidimensionaler rekursiver Filter	107
5.1 Ortsbereichsentwurfsverfahren	107

5.2 Entwurf im Frequenzbereich	108
5.3 Neue Verfahren zum Entwurf zweidimensio- naler rekursiver stabiler Filter	110
5.3.1 Entwurfsverfahren 1	111
5.3.1.1 Transformation eines Blocks 2.Grades	111
5.3.1.2 Untersuchung der Stabilität	117
5.3.1.3 Transformation eines Blocks 2.Grades mit komplexen Trans- formationsparametern	118
5.3.1.4 Stabilitätsbedingungen für die Transformationsparameter	120
5.3.1.5 Beispiele	125
5.3.1.5.1 Beispiel 1: Potenztiefpaß	126
5.3.1.5.2 Beispiel 2: Elliptisches Filter	130
5.3.2 Entwurfsverfahren 2	134
5.3.2.1 Drehung eines eindimensio- nalen Filters	134
5.3.2.2 Kaskadierung von um $\nu \cdot \pi/4$ gedrehten 1-D Filtern	139
5.3.2.3 Einschränkung der Sperrgrenze	140
5.3.2.4 Anforderungen an die Toleranzen der 1-D-Filter	142
5.3.2.5 Verallgemeinerung	146
5.3.2.6 Beispiele	149
5.3.2.6.1 Rotationssymmetrischer Tiefpaß	149
5.3.2.6.2 Rotationssymmetrischer Tiefpaß mit Nullphase	151
5.3.3 Einfacher Entwurf von Fan-Filtern	153
5.3.3.1 Entwurfsprinzip	154
5.3.3.2 Beispiel	164
5.3.4 Entwurfsverfahren 3	165
5.3.4.1 Prinzip	168
5.3.4.2 Beispiel	173

6. Anwendungsbeispiele für 2-D rekursive digitale Filter	175
6.1 Rauschunterdrückung bei der automatischen Auswertung von Mikro-Chromatogrammen	175
6.2 Rekursive Hochpaßfilterung zur Hervorhebung von Bilddetails	179
6.3 Anwendung eines rekursiven Kantendetektionsfilters	179
7. Zusammenfassung	185
Anhang 1	186
Anhang 2	189
Literaturverzeichnis	192

Liste der wichtigsten verwendeten Formelzeichen

$B_v(z), B_v(z_\mu)$	- Übertragungsfunktion eindimensionaler diskreter nichtrekursiver Teilsysteme Zählerpolynom eines 1-D-Systems
$B(z_1, z_2)$	- Zählerpolynom eines zweidimensionalen rekursiven Systems
$\underline{B}, \underline{B}_v$	- Zählerkoeffizientenmatrix
$b_\mu, \hat{b}_\mu$	- Zählerkoeffizient eines eindimensionalen Systems
$b_{\mu v}$	- Zählerkoeffizient eines zweidimensionalen Systems
$\tilde{b}_{\mu v}$	- Zählerkoeffizient eines ortsvarianten Systems
$C_v(z), C_v(z_\mu)$	- Nennerpolynom eines eindimensionalen diskreten Systems
$C(z_1, z_2), C_v(z_1, z_2)$	- Nennerpolynom eines zweidimensionalen diskreten Systems
$\underline{C}, \underline{C}_v$	- Matrix der Nennerkoeffizienten
$c_\mu, \hat{c}_\mu$	- Nennerkoeffizient eines eindimensionalen Systems
$c_{\mu v}$	- Nennerkoeffizient eines zweidimensionalen Systems
$\mathbb{C}$	- Menge der komplexen Zahlen
$DFT\{\dots\}$	- Diskrete Fouriertransformierte (DFT) von ...
$DFT^{-1}\{\dots\}$	- Inverse DFT von ...

d_0, d_1, d_2, d_3	- Koeffizienten kontinuierlicher Systeme
f	- Frequenz
f_A	- Abtastfrequenz
$G(w_1, w_2)$	- Teilübertragungsfunktion eines kontinuierlichen Systems
$H(z), H_\lambda(z)$	- Übertragungsfunktionen eindimensionaler digitaler Systeme
$\ddot{H}_{\dots}(z_1, z_2)$	- Übertragungsfunktionen eines zweidimensionalen diskreten Systems ... = Indizes bzw. Zeichen zur Unterscheidung
$h_0(k, l)$	- Impulsantwort eines zweidimensionalen diskreten Systems
k, k_1	- diskrete Ortsvariable
l, l_1	- diskrete Ortsvariable
L_∞, L_2	- Fehlnormen
n	- Filtergrad
N	- Menge der natürlichen Zahlen
N	- Maß für die Dimension eines Bildes
N_B	- Maß für den verfügbaren Datenspeicherplatz im Hauptspeicher (Zahl der reellen Worte)
r_v	- ganzzahlige Indizes

$\mathbb{R}$	- Menge der reellen Zahlen
$S\{\dots\}$	- Systemreaktion auf ...
$\underline{T}_1, \underline{T}_2$	- Transformationsmatrizen
$T_{acc}, T_{add}, T_{mult}, T_{tran}$	- Operationszeiten
t_A	- Zeit für die arithmetische Berechnung
t_{IO}	- Zeit für die Datenein- und ausgabe
t_G	- Gesamtrechnzeit
$t_{\mu v}$	- komplexer Transformationsparameter
$U(z_1, z_2)$	- zweidimensionale Z-Transformierte des Eingangssignals
$U_1(z_1)$	- Z-Transformierte der l-ten Zeile des Eingangssignals
$u(k, l)$	- zweidimensionales diskretes Eingangssignal
$\underline{u}_l$	- Vektor der l-ten Zeile des Eingangsbildes
$V(\kappa, \lambda), V_C(\kappa, \mu)$	- diskrete Fouriertransformierte
$v_1(k, l), v_2(k, l), v_3(k, l)$	- erweiterte zweidimensionale periodische diskrete Folgen
w_M	- Faltungskern der DFT: $e^{-j\frac{2\pi}{M}}$
w	- auf die Abtastfrequenz normierte komplexe Kreisfrequenz kontinuierlicher Systeme
$x(k, l)$	- allg. diskretes Signal

x	- freier Optimierungsparameter
$Y(z_1, z_2)$	- zweidimensionale Z-Transformierte des Eingangssignals
$Y_1(z_1)$	- Z-Transformierte der 1-ten Zeile des Ausgangssignals
$y(k, l)$	- zweidimensionales diskretes Ausgangssignal
$\underline{y}_1$	- Vektor der 1-ten Zeile des Ausgangssignals
$\mathbb{Z}$	- Menge der ganzen Zahlen
$Z\{\dots\}$	- Z-Transformation von ...
$Z^{-1}\{\dots\}$	- Inverse Z-Transformation von ...
z, z_1, z_2, z_v	- komplexe Frequenzvariable
z_{0v}	- Nullstelle der Übertragungsfunktion
$z_{\infty v}$	- Polstelle der Übertragungsfunktion
$\underline{z}_v$	- Vektoren, deren Elemente Potenzen von z_v enthalten
$\underline{z}_v^T$	- transponierter Vektor $\underline{z}_v$
α	- Transformationsparameter
β	- Transformationsparameter oder Drehwinkel
β_0	- Winkel
$\beta_{\mu v}$	- Zählerkoeffizienten 2-D kontinuierlicher Filter
γ_λ, γ	- komplexe Transformationsparameter

$\delta_1, \delta_1^{(\dots)}$	- zul. Toleranzen im Durchlaßbereich
$\delta_2, \delta_2^{(\dots)}$	- zul. Toleranzen im Sperrbereich
Δ	- Approximationsfehler
ϵ	- Fehlerschranke, Genauigkeits- schranke, Stabilitätsschranke
$\zeta_{\mu\nu}$	- Nennerkoeffizienten 2-D kontinu- ierlicher Filter
$\kappa, \lambda, \mu, \nu, \rho$	- ganzzahlige Laufvariable
λ_1, ρ_1	- reelle Transformationsparameter
ϕ	- Optimierungsparameter
φ	- Winkel
$\Omega, \Omega_1, \Omega_2$	- (normierte) reelle Frequenzvariable ($\Omega=\pi=180^\circ$ entspricht der halben Abtastfrequenz)
$\Omega_D'S$	- Durchlaß- bzw. Sperrgrenze selektiver Filter

1. Einführung und Übersicht

Die Verarbeitung zweidimensionaler digitaler Signale gewinnt für viele Anwendungsbereiche zunehmend an Bedeutung. Seit etwa zehn Jahren ist die Bildverarbeitung, das wichtigste Gebiet der zweidimensionalen Signalverarbeitung, ein Hauptgegenstand der Forschung. Gerade in den letzten Jahren wurden dafür immer neue Anwendungsmöglichkeiten erschlossen.

Methoden der Bildverbesserung, Bildsegmentation, Merkmalsextraktion bzw. Bildanalyse finden heute Anwendung in

- Medizin (Mikro-Chromatographie, Tomographie, Zell-, Blutkörper-, Chromosomen-Analyse, Röntgentechnik, Ultraschallaufnahmetechnik, Szintigraphie, Knochenvermessung nach Hüftgelenkoperationen in der Orthopädie),
- Biologie (Beobachtung von Wachstumsprozessen),
- Geophysik (Magnet- und Gravitationsfelder, Seismische Signale, Auswertung von Luftbildern, Vermessung des Meeresbodens),
- Meteorologie (Temperatur- und Luftdruckmessungen),
- Astronomie (Trennung von dicht beieinander liegenden Gestirnen),
- Archäologie (Luft- und Röntgenbilddauswertung),
- Physik (Analyse von Teilchenspuren, Elektronenmikroskopie),
- Kriminalistik (Analyse von Fingerabdrücken, Gesichtsprofilen, Geschoßspuren),

- Militärtechnik (Radar- und Sonarbildverarbeitung,
Luftbildauswertung),
- Satellitentechnik (Verbesserung der Funkbilder von
Forschungssatelliten),
- Industrie (Prozeßautomation, zerstörungsfreie
Werkstoffprüfung).

Dementsprechend groß ist die Zahl der bisher erschienenen Veröffentlichungen. Andererseits lassen sich aus den vielen verschiedenen Anwendungen einige wichtige und immer wiederkehrende Aufgabenstellungen extrahieren. Dazu gehört die lineare zweidimensionale digitale Filterung, die Gegenstand dieser Arbeit ist.

Im 2.Kapitel wird zunächst eine Übersicht über die Systemeigenschaften und die Beschreibungsmöglichkeiten zweidimensionaler digitaler Systeme gegeben.

Das 3.Kapitel befaßt sich mit zweidimensionalen nichtrekursiven Filtern. Dabei werden einige bisher in der Fachliteratur nicht oder nur oberflächlich behandelte Probleme genauer untersucht. Die Verarbeitung der zweidimensionalen digitalen Signale erfolgt heute oftmals an einem Minicomputer. Wegen der großen zu verarbeitenden Datenmengen muß, anders als bei der eindimensionalen Signalverarbeitung, der Aufwand für den Datentransfer bei der Beurteilung von Filteralgorithmen mit berücksichtigt werden. Beispielhaft für ähnliche Rechnerkonfigurationen wurde für eine PDP 11/45 eine genaue Analyse des Rechenaufwands bei der Realisierung der zweidimensionalen nichtrekursiven Filterung mit und ohne schnelle Faltung durchgeführt. Für die Ausführung der schnellen Faltung mit Hilfe der zweidimensionalen FFT wurde ein eigener Algorithmus zur Matrixtransponierung mit gleichzeitigem Bitreversal entwickelt.

Das 4.Kapitel befaßt sich mit der Beschreibung und der Realisierung zweidimensionaler rekursiver Filter. Zunächst wer-

den einige für rekursive Filter wichtige Begriffe und Eigenschaften definiert. Ein Hauptproblem bei zweidimensionalen rekursiven Filtern ist die Prüfung der Stabilität. Es wird ein neuer numerischer Stabilitätstest vorgestellt, der gegenüber bisher bekannten Verfahren [48,49] wegen seiner Einfachheit deutliche Rechenzeitvorteile bringt. Für zwei verschiedene Beschreibungsmöglichkeiten von zweidimensionalen rekursiven Digitalfiltern werden die Unterschiede hinsichtlich des Realisierungsaufwands besprochen. Den Abschluß dieses Kapitels bildet ein Vergleich des Realisierungsaufwands von nichtrekursiver und rekursiver Filterung bei Verwendung eines Minicomputers.

Im nachfolgenden 5.Kapitel wird der Entwurf von rekursiven zweidimensionalen digitalen Filtern behandelt. Dazu werden vier neue Entwurfsmethoden vorgestellt, die gegenüber bekannten Verfahren deutliche Vorteile hinsichtlich des Entwurfs- und Realisierungsaufwands besitzen. Diese Verfahren gehen von einem eindimensionalen stabilen rekursiven Prototypfilter aus. Dadurch wird der Optimierungsaufwand erheblich reduziert bzw. er entfällt völlig (beim Verfahren 2). Gleichzeitig ist bei dieser Methode von vornherein garantiert, daß die damit entworfenen zweidimensionalen Filter stets stabil sind, obgleich die zur Optimierung verwendeten Parameter nicht eingeschränkt werden.

Im 6.Kapitel werden anhand von drei Anwendungsbeispielen Einsatzmöglichkeiten für zweidimensionale rekursive digitale Filter aufgezeigt.

2. Lineare digitale Systeme zur 2-D Signalverarbeitung

Die Theorie zweidimensionaler Systeme zur Verarbeitung diskreter zweidimensionaler (2-D) Signale stellt im wesentlichen eine Erweiterung der Theorie eindimensionaler digitaler Systeme dar [1].

2.1 Zweidimensionale diskrete Signale

Ein zweidimensionales diskretes Signal ist eine Folge $\{x(k,l)\}$ von reellen oder komplexen Werten, die von zwei orts- und/oder zeitdiskreten Variablen k,l abhängt.

Aus Realisierbarkeitsgründen muß dabei mindestens eine der Variablen endlich sein. Die Folgen erhält man im allgemeinen durch Abtastung eines kontinuierlichen zweidimensionalen Signals. Im weiteren werden nur Folgen betrachtet, deren Variablen k und l so normiert sind, daß sie ein äquidistantes Gitter bilden und $k,l \in \mathbb{Z}$ ($\mathbb{Z}$ = Menge der ganzen Zahlen) ist. Statt $\{x(k,l)\}$ wird vereinfachend häufig $x(k,l)$ geschrieben.

Eine zweidimensionale diskrete Folge $x(k,l)$ heißt *k a u - s a l*, wenn sie ausschließlich im 1.Quadranten der (k,l) -Ebene nichtverschwindende Werte besitzt (Bild 2.1).

Bei *o n - l i n e - V e r a r b e i t u n g* ist mindestens eine Variable physikalisch eine diskrete Zeitvariable. Bezüglich dieser Variablen muß das Signal kausal, aber nicht endlich sein.

Bei Verarbeitung von 2-D diskreten Signalen in einem Mini-computer liegt im allgemeinen *o f f - l i n e - V e r a r b e i t u n g* vor. In diesem Fall muß das gesamte Signal abgespeichert werden. Dies impliziert stets endliche Signale in beiden Variablen k,l . Die Forderung der Kausalität ist dann nicht mehr notwendig.

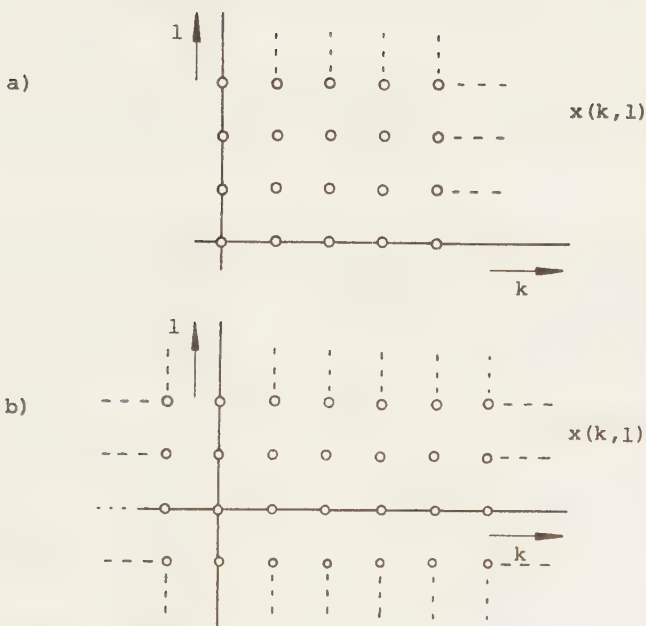


Bild 2.1: Beispiel einer
 a) kausalen
 b) nichtkausalen Folge

Bei der Bildverarbeitung, dem Hauptanwendungsgebiet der zweidimensionalen Signalverarbeitung, liegen stets endliche Eingangsfolgen vor. Durch Verschiebung kann deshalb erreicht werden, daß das Eingangssignal kausal wird.

Im weiteren werden

- kausale endliche Eingangssignale und
- off-line-Verarbeitung vorausgesetzt.

2.2 Definition und Eigenschaften zweidimensionaler Systeme | 2-4 |

Ein allgemeines zweidimensionales (2-D) System S (Bild 2.2) bildet eine Eingangsfolge $u(k,l)$

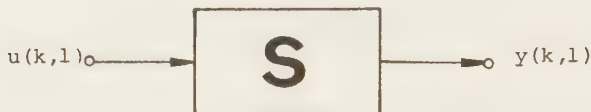


Bild 2.2 Allgemeines 2-D System

eindeutig auf eine Ausgangsfolge

$$y(k,l) = S\{u(k,l)\} \quad k,l \in \mathbb{Z} \quad (2.2-1)$$

ab, wobei Ein- und/oder Ausgangssignal reell- oder komplexwertig sind.

Im folgenden werden nur reellwertige Systeme betrachtet. Sie bilden reelle Eingangssignale stets auf reellwertige Ausgangssignale ab.

Besondere Bedeutung haben lineare Systeme. Das System S ist genau dann linear, wenn bei jeder Linearkombination von zwei Eingangssignalen $u_1(k,l)$, $u_2(k,l)$ die Systemreaktion auf das Summensignal darstellbar ist als dieselbe Linearkombination der Systemreaktion auf die Einzelsignale, d.h., wenn

$$S\{a_1 u_1(k,l) + a_2 u_2(k,l)\} = a_1 S\{u_1(k,l)\} + a_2 S\{u_2(k,l)\} \quad (2.2-2)$$

$$a_1, a_2 \in \mathbb{C} \quad , \quad u_1, u_2 \in F$$

$\mathbb{C}$ = Menge der komplexen Zahlen

F = Menge aller Folgen $\{u(k,l)\}$ gilt.

Im weiteren werden im allgemeinen v e r s c h i e b u n g s i n -
v a r i a n t e S y s t e m e betrachtet. Ist

$$y(k,l) = S\{u(k,l)\}$$

$$\text{und } y(k-k_0, l-l_0) = S\{u(k-k_0, l-l_0)\} \quad (2.2-3)$$

$$k_0, l_0 \in \mathbb{Z},$$

so heit das System verschiebungsinvariant.

Die Systemreaktion ist dann invariant gegen eine Verschiebung der Erregung. Da bei 2-D-Signalen den Variablen k, l im allgemeinen Ortsvariable zugeordnet werden, spricht man auch von o r t s i n v a r i a n t e n S y s t e m e n.

Ein Spezialfall sind s e p a r a b l e S y s t e m e.

$$\text{Sei } u(k,l) = u_1(k) \cdot u_2(l),$$

$$\begin{aligned} \text{wenn } y(k,l) &= S\{u(k,l)\} = \\ &= S\{u_1(k) \cdot u_2(l)\} = \\ &= S\{u_1(k)\} \cdot S\{u_2(l)\} = \\ &= y_1(k) \cdot y_2(l) \end{aligned} \quad (2.2-4)$$

gilt, ist das System S separabel.

Von Bedeutung sind ferner s t a b i l e S y s t e m e.

Ein System ist stabil, wenn es auf beschrnkte Eingangssignale mit beschrnkten Ausgangssignalen reagiert.

Mathematisch formuliert bedeutet dies, mit

$$L_\infty = \{v \in F : \|v\|_\infty < \infty\}$$

$$\text{und } \|u\|_\infty = \sup_{k,l \in \mathbb{Z}} |u(k,l)| \quad (2.2-5)$$

ist das System S genau dann stabil, wenn für alle $u \in L_\infty$

$$\text{mit } y(k,l) = S\{u(k,l)\} \quad (2.2-6)$$

$$y \in L_\infty \quad (2.2-7)$$

gilt.

In der englischsprachigen Literatur spricht man dann häufig von BIBO-Stabilität (bounded intput - bounded output).

2.3 Beschreibung zweidimensionaler linearer Systeme

Die Beschreibung linearer diskreter Systeme erfolgt wie üblich im Zeit- oder Frequenzbereich. Den Zusammenhang beider Bereiche liefert die Z-Transformation. Sie ordnet einer zeitlichen Folge eindeutig eine Funktion aus dem Frequenzbereich zu und umgekehrt. Da bei zweidimensionalen Signalen $u(k,l)$ die Variablen $k, l \in \mathbb{Z}$ im allgemeinen Ortsvariablen darstellen, spricht man anstelle vom Zeitbereich vom Ortsbereich.

Die Beschreibung im Ortsbereich erfolgt wahlweise durch

- Impulsantwort
- Zustandsgleichungen
- Differenzengleichungen.

Die Beschreibung im Frequenzbereich erfolgt durch die Übertragungsfunktion.

2.3.1 Beschreibung durch die Impulsantwort [3]

Jede beliebige diskrete Eingangsfolge $u(k,l)$ lässt sich durch

$$u(k,l) = \sum_{\mu=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} u(\mu,\nu) \cdot \gamma_O(k-\mu, l-\nu) \quad (2.3-1)$$

beschreiben, wobei

$$\gamma_O(k,l) = \begin{cases} 1 & k = l = 0 \\ 0 & \text{sonst} \end{cases} \quad (2.3-2)$$

den Einheitsimpuls darstellt.

Die I m p u l s a n t w o r t $h_O(k,l)$ ist die Systemreaktion auf den Einheitsimpuls $\gamma_O(k,l)$

$$y(k,l) = h_O(k,l) = S\{\gamma_O(k,l)\}. \quad (2.3-3)$$

Mit der Reaktion auf einen um μ, ν Werte verschobenen Einheitsimpuls

$$h_O(k,\mu,l,\nu) = S\{\gamma_O(k-\mu, l-\nu)\} \quad (2.3-4)$$

ergibt sich wegen der Linearität des Systems nach (2.2-2) mit (2.3-1)

$$y(k,l) = \sum_{\mu=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} u(\mu,\nu) \cdot h_O(k,\mu,l,\nu). \quad (2.3-5)$$

Bei v e r s c h i e b u n g s i n v a r i a n t e n Systemen gilt außerdem nach (2.2-3)

$$h_O(k,\mu,l,\nu) = h_O(k-\mu, l-\nu) = S\{\gamma_O(k-\mu, l-\nu)\}. \quad (2.3-6)$$

Damit geht (2.3-5) über in

$$y(k,l) = \sum_{\mu=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} u(\mu,\nu) \cdot h_0(k-\mu,l-\nu) . \quad (2.3-7)$$

Die Summe (2.3-7) stellt eine zweidimensionale Faltung des Eingangssignals $u(k,l)$ mit der Impulsantwort $h_0(k,l)$ dar. Abgekürzt schreibt man

$$y(k,l) = u(k,l) ** h_0(k,l) . \quad (2.3-8)$$

Für ein 2-D lineares verschiebungsinvariantes System genügt demnach die Impulsantwort $h_0(k,l)$, um das System vollständig zu beschreiben.

Ein derartiges System heißt *k a u s a l*, wenn die Impulsantwort $h_0(k,l)$ kausal ist, d.h. wenn sie nur im 1.Quadranten der (k,l) -Ebene nichtverschwindende Werte besitzt.

Ein System ist *s t a b i l*, wenn dessen Impulsantwort absolut summierbar ist, wenn also

$$\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} |h_0(k,l)| < M < \infty , \quad (2.3-9)$$

gilt, wobei M bei der praktischen Realisierung durch den Zahlenbereich des Systems festgelegt ist.

Da bei rekursiven Systemen die Impulsantwort unendlich ausgedehnt ist, bedient man sich dort zweckmäßigerweise anstelle der Beschreibung durch die Impulsantwort der Beschreibung durch die Übertragungsfunktion oder die Differenzengleichung.

2.3.2 Beschreibung durch die Übertragungsfunktion |3|

Durch Anwendung der zweidimensionalen Z-Transformation |56| auf (2.3-7) erhält man

$$Z\{y(k,l)\} = Y(z_1, z_2) = U(z_1, z_2) \cdot \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h_0(k,l) z_1^{-k} z_2^{-l}, \quad (2.3-10)$$

wobei

$$H(z_1, z_2) = Z\{h_0(k,l)\} = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h_0(k,l) z_1^{-k} z_2^{-l} \quad (2.3-11)$$

als Übertragungsfunktion bezeichnet wird.

$$\text{Mit} \quad Y(z_1, z_2) = H(z_1, z_2) \cdot U(z_1, z_2) \quad (2.3-12)$$

wird die zweidimensionale Faltung im Ortsbereich nach (2.3-8) durch die Z-Transformation in eine Multiplikation im z-Bereich übergeführt. Für

$$z_v = e^{j\Omega_v}, \quad \Omega_v = 2\pi \cdot \frac{f_v}{f_{Av}}, \quad f_{Av} = \text{Abtastfrequenz } v \in \{1, 2\},$$

erhält man die Übertragungsfunktion im Frequenzbereich (f_1, f_2) . Voraussetzung für (2.3-12) ist allerdings, daß die Konvergenzbereiche von $U(z_1, z_2)$ und $H(z_1, z_2)$ überlappen.

Eine für die Realisierung wichtige Klasse von Übertragungsfunktionen ist analog zum eindimensionalen, die Klasse der rationalen Funktionen. Dann ist die Übertragungsfunktion nach (2.3-11) darstellbar als

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)} = \frac{\sum_{\mu} \sum_{\nu} b_{\mu\nu} z_1^{-\mu} z_2^{-\nu}}{\sum_{\mu} \sum_{\nu} c_{\mu\nu} z_1^{-\mu} z_2^{-\nu}} \quad (2.3-13)$$

Derartige Systeme lassen sich stets durch eine Differenzengleichung endlicher Ordnung beschreiben. Im weiteren werden wir uns auf diese Systeme beschränken, welche wir auch als *digitale Filter* bezeichnen. Dabei sind die Polynome $B(z_1, z_2)$, $C(z_1, z_2)$ in (2.3-13) endlich, d.h.

$$B(z_1, z_2) = \sum_{\mu=-r_1}^{m_1} \sum_{\nu=-r_2}^{m_2} b_{\mu\nu} z_1^{-\mu} z_2^{-\nu} \quad (2.3-14a)$$

und

$$C(z_1, z_2) = \sum_{\mu=-q_1}^{n_1} \sum_{\nu=-q_2}^{n_2} c_{\mu\nu} z_1^{-\mu} z_2^{-\nu} \quad (2.3-14b)$$

mit $m_1, m_2, n_1, n_2, r_1, r_2, q_1, q_2 \in \mathbf{N}$ ($\mathbf{N}$ = Menge der natürlichen Zahlen).

2.3.3 Beschreibung durch eine Differenzengleichung

Für ein allgemeines lineares verschiebungsinvariantes System mit rationaler Übertragungsfunktion ergibt sich aus (2.3-12... 2.3-14b) die partielle Differenzengleichung

$$\sum_{\mu=-q_1}^{n_1} \sum_{\nu=-q_2}^{n_2} c_{\mu\nu} y(k-\mu, l-\nu) = \sum_{\mu=-r_1}^{m_1} \sum_{\nu=-r_2}^{m_2} b_{\mu\nu} u(k-\mu, l-\nu) \quad (2.3-15)$$

Für $c_{\mu\nu} = 0$ falls $(\mu, \nu) \neq (0, 0)$ heißt das System *n i c h t r e k u r s i v*, andernfalls *r e k u r s i v*, obwohl die Differenzengleichung (2.3-15) in dieser allgemeinen Form nur dann rekursiv berechenbar ist, wenn bestimmte Kausalitätsbedingungen erfüllt sind (siehe Kap.4.1.3).

Neben den hier angegebenen Beschreibungsmöglichkeiten lassen sich lineare Systeme noch durch Zustandsgleichungen beschreiben [1,2,4].

Bisher wurde stets angenommen, daß das betrachtete System einen Eingang und einen Ausgang hat. Diese Beschreibung wurde gewählt, weil sie auf die Charakterisierung des Systems im Orts- und Frequenzbereich durch je eine Funktion (Impulsantwort, Übertragungsfunktion) führt. Man gewinnt damit eine relativ einfache Vorschrift für den Entwurf derartiger Systeme. Eine Einschränkung über die Ausdehnung des Ein- und Ausgangssignals wird dabei nicht gemacht.

Bei der Verarbeitung müssen die 2-D Signale schon aus Realisierungsgründen zumindest längs einer Ortsvariablen auf einen endlichen Bereich beschränkt sein.

Die Verarbeitung erfolgt dann entweder zeilen- oder spaltenweise, wobei die Zeilen oder Spalten endliche Dimension N besitzen. Daher kommt eine andere Form der Beschreibung der Realisierung näher, die das 2-D-System als ein eindimensionales System mit N Ein- und Ausgängen beschreibt [2].

Bei der Bildverarbeitung, dem Hauptanwendungsgebiet der 2-D Signalverarbeitung, beschränkt man sich für das Ein- und Ausgangssignal im allgemeinen auf einen gleichgroßen Ausschnitt. Eine systemtheoretische Beschreibung für ein derartiges System mit $N_1 \times N_2$ Ein- und Ausgängen wurde erstmals in [4] vorgestellt.

Im weiteren beschränken wir uns zunächst auf die Beschreibung durch ein System mit einem Ein- und Ausgang. Erst bei der Realisierung wird versucht, den Zusammenhang bzw. die

Unterschiede zu den eben erwähnten Beschreibungsformen anzugeben.

In den nachfolgenden beiden Kapiteln soll genauer auf bestimmte Eigenschaften und spezielle Gegebenheiten bei der Implementierung von nichtrekursiven und rekursiven 2-D Filtern an einem Minicomputer PDP 11/45 eingegangen werden. Dabei sollen einige Punkte genauer behandelt und neue Ergebnisse vorgestellt werden.

Zur Vereinfachung der Darstellung werden dabei ideale Systeme in der bisher definierten Weise angenommen. Es wird also vorausgesetzt, daß

- Koeffizienten und Zustandsvariablen reell sind,
- alle arithmetischen Operationen exakt ausgeführt werden.

Das bedeutet, daß die durch endliche Wortlänge verursachten Effekte |9-12| hier nicht betrachtet werden. Unter diesen Voraussetzungen folgt:

- alle unterschiedlichen Realisierungen eines Systems verhalten sich exakt gleich.

3. Zweidimensionale nichtrekursive Filter

3.1 Beschreibung durch Differenzengleichung und Übertragungsfunktion

Mit $c_{00} = 1$ und $c_{\mu\nu} = 0$ für $(\mu, \nu) \neq (0, 0)$ ergibt sich aus (2.3-15) die Differenzengleichung eines linearen verschiebungsinvarianten 2-D nichtrekursiven Filters

$$y(k, l) = \sum_{\mu=-r_1}^{m_1} \sum_{\nu=-r_2}^{m_2} b_{\mu\nu} u(k-\mu, l-\nu). \quad (3.1-1)$$

Die zugehörige Übertragungsfunktion erhält man durch Z-Transformation zu

$$H(z_1, z_2) = \frac{Y(z_1, z_2)}{U(z_1, z_2)} = \sum_{\mu=-r_1}^{m_1} \sum_{\nu=-r_2}^{m_2} b_{\mu\nu} z_1^{-\mu} z_2^{-\nu}. \quad (3.1-2)$$

$$\text{Mit } \underline{z}_\rho^T = [z_\rho^r, z_\rho^{r-1}, \dots, 1, z_\rho^{-1}, \dots, z_\rho^{-m_\rho}] \quad (3.1-3)$$

$$\rho \in \{1, 2\}$$

und der Koeffizientenmatrix

$$\underline{B} = [b_{\mu\nu}] = \begin{bmatrix} b_{-r_1, -r_2} & b_{-r_1, -r_2+1} & \dots & b_{-r_1, 0} & \dots & b_{-r_1, m_2} \\ b_{-r_1+1, -r_2} & b_{-r_1+1, -r_2+1} & \dots & b_{-r_1+1, 0} & \dots & b_{-r_1+1, m_2} \\ \vdots & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{0, -r_2} & b_{0, -r_2+1} & \dots & b_{0, 0} & \dots & b_{0, m_2} \\ \vdots & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{m_1, -r_2} & b_{m_1, -r_2+1} & \dots & b_{m_1, 0} & \dots & b_{m_1, m_2} \end{bmatrix}$$

ergibt sich die zu (3.1-2) äquivalente Matrixschreibweise

$$H(z_1, z_2) = \underline{z}_1^T \underline{B} \underline{z}_2 . \quad (3.1-5)$$

Durch Koeffizientenvergleich von (2.3-11) und (3.1-2) folgt

$$b_{\mu\nu} = h_0(\mu, \nu),$$

$$\text{d.h. } h_0(k, l) = 0 \quad \forall (k, l) \notin \{-r_1, m_1; -r_2, m_2\}.$$

Die Impulsantwort eines nichtrekursiven Systems ist demnach endlich und umfaßt genau den Ausdehnungsbereich der Koeffizientenmaske $\underline{B}$ (3.1-4). Daher werden derartige Systeme auch als FIR-Filter (finite impulse response)-Filter bezeichnet.

Für $r_1 = r_2 = 0$ nimmt die Impulsantwort $h_0(k, l)$ ausschließlich im 1. Quadranten der (k, l) -Ebene nichtverschwindende Werte an, das Filter ist in diesem Fall kausal. Durch Verschiebung der Koeffizientenmaske, d.h. Multiplikation der Übertragungsfunktion nach (3.1-2) mit $z_1^{-r_1} \cdot z_2^{-r_2}$ kann jedes nichtkausale nichtrekursive Filter in ein kausales Filter mit gleichem Betragsfrequenzgang übergeführt werden [1].

Nichtrekursive Systeme besitzen zwei Eigenschaften, die sie besonders auszeichnen:

- sie sind stets stabil,
- sie können strikt linearphasig sein [1].

Deshalb werden nichtrekursive 2-D Filter seit langem in der 2-D Signalverarbeitung angewandt. Der Entwurf derartiger Filter ist in der Literatur bereits eingehend untersucht worden [5, 6, 8, 13-17]. Darauf wird deshalb nicht weiter eingegangen.

Im folgenden werden einige Probleme behandelt, die generell bei der Verarbeitung mit nichtrekursiven Systemen von Bedeutung sind.

3.2 _Spezielle_Probleme

3.2.1 Vergrößerung des Ausgangsbildes, Randeffekte

Bei der digitalen Bildverarbeitung liegt in den meisten Fällen das Eingangssignal $u(k,l)$ in Form eines endlichen Bildes vor.

Deshalb sei $u(k,l) = 0$ für $(k,l) \notin \{0,K;0,L\}$

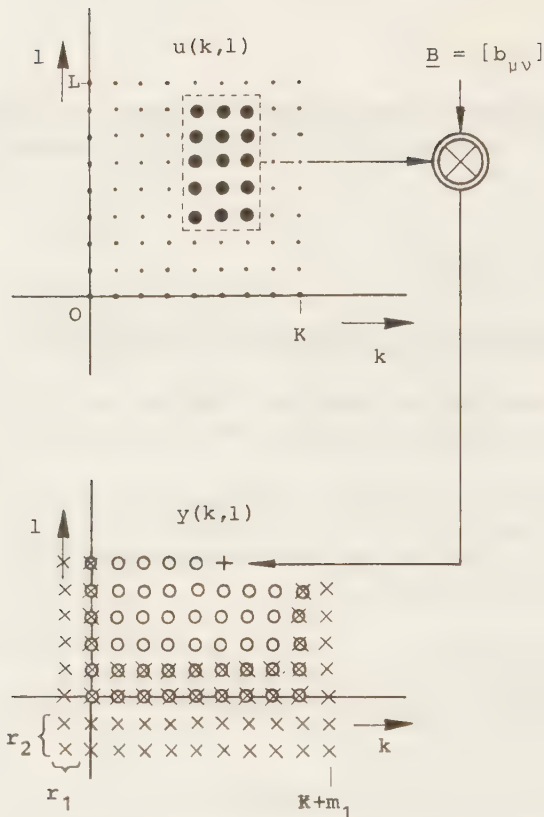
Die Verarbeitung erfolge zeilenweise, beginnend von der linken unteren Ecke (Bild 3.1).

Als Beispiel sei in (3.1-1) speziell $m_1=1, r_1=1, m_2=2, r_2=2$ gewählt. Da in diesem Fall $r_v \neq 0 \forall v \in \{1,2\}$ gilt, ist das Filter in beiden Richtungen nichtkausal. Es ergibt sich

$$y(k,l) = \sum_{\mu=-1}^1 \sum_{v=-2}^2 b_{\mu v} u(k-\mu, l-v). \quad (3.2-1)$$

Aus Bild 3.1 sind folgende Eigenschaften erkennbar:

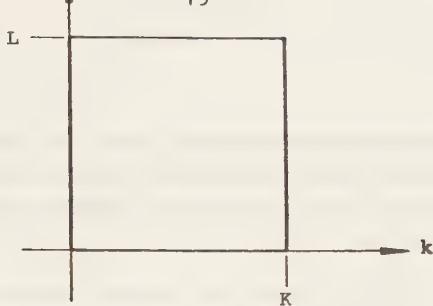
1. Das Ausgangsbild vergrößert sich gegenüber dem Eingangsbild um einen Rand. Dabei umfaßt der
 - obere Rand m_2 Punkte, der
 - untere Rand r_2 Punkte, der
 - linke Rand r_1 Punkte und der
 - rechte Rand m_1 Punkte (Bild 3.2b).
2. Die $m_v + r_v, v \in \{1,2\}$ Punkte an den Rändern des Ausgangsbildes beinhalten einen Ein- oder Ausschwinganteil oder beides (Bild 3.2b).



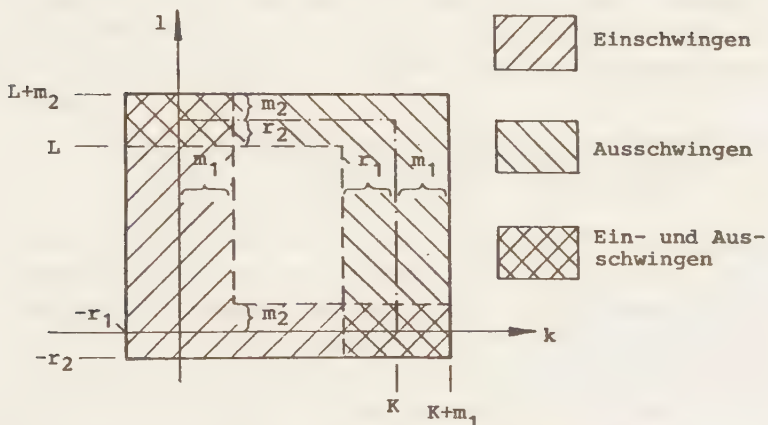
- × Punkte des Ausgangsbildes, die außerhalb des ursprünglichen Bildbereichs liegen.
- Ø Punkte des Ausgangsbildes, die Ein-/bzw. Ausschwinganteile enthalten.
- ×, Ø, Ø Bereits berechnete Bildpunkte des Ausgangsbildes
- + für die angegebene Stellung der Maske berechneter Ausgangswert

Bild 3.1: Zeilenweise Verarbeitung mit einem nichtkausalen nichtrekursiven System nach Gl. (3.2-1).
Berechnung von $y(5,5)$.

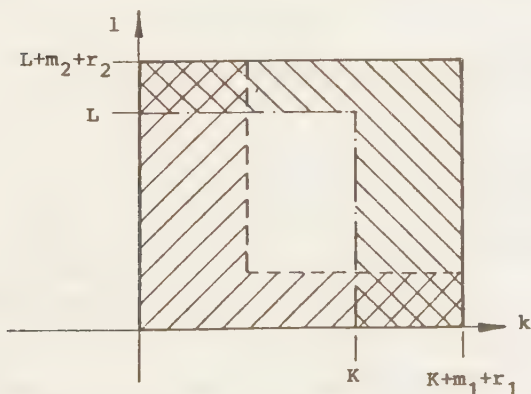
3. Im Falle der Verarbeitung mit einem äquivalenten kausalen System vom Grad (m_1+r_1, m_2+r_2) ergibt sich noch eine zusätzliche Verschiebung um (r_1, r_2) Werte in Verarbeitungsrichtung (Bild 3.2c).



a) Bildbereich des Eingangsbildes $u(k,l)$



b) Bildbereich des mit einem nichtkausalen FIR-Filter gefilterten Ausgangsbildes $y(k,l)$



c) Bildbereich des mit einem kausalen FIR-Filter gefilterten Ausgangsbildes $y(k,l)$

Bild 3.2: Nichtrekursive 2-D Filterung mit kausalem und nichtkausalem System

Im allgemeinen beschränkt man sich bei der digitalen Verarbeitung von Bildern beim Ausgangsbild $y(k,l)$ auf einen gleichgroßen Ausschnitt von $(K+1) \cdot (L+1)$ Punkten wie für das vorliegende Eingangsbild $u(k,l)$.

Zur Vermeidung bzw. Verminderung der Ränder, verursacht durch Ein- bzw. Ausschwingvorgänge, kann man entweder mit einem ortsvarianten System arbeiten [4], oder zunächst in geeigneter Weise eine Extrapolation des Eingangsbildes über seine Ränder hinaus vornehmen und das so erweiterte Bild ortsinvariant verarbeiten.

3.2.2 Ortsvariante Filterung zur Vermeidung des Randeffektes

Betrachtet sei stets die untere linke Ecke und der linke Rand des Eingangs- und Ausgangsbildes (Bild 3.2). Dies bedeutet keine Einschränkung, da nachfolgende Überlegungen sinngemäß für alle Ecken und Ränder gelten. Ferner sei $(k,l) \in \{0,K;0,L\}$. Gefordert wird, daß die Systemreaktion $\tilde{y}(k,l)$ des ortsvarianten Filters

$$\tilde{y}(k,l) = \sum_{\mu=-r_1}^{\alpha} \sum_{v=-r_2}^{\beta} \tilde{b}_{\mu v}(k,l) u(k-\mu, l-v) \quad (3.2-2)$$

mit den ortsvarianten Koeffizienten $\tilde{b}_{\mu v}(k,l)$

$$\alpha = \text{Min } \{m_1, k\}$$

$$\beta = \text{Min } \{m_2, l\}$$

in den Randpunkten

$$(k,l) \in \begin{cases} \{0, m_1-1; 0, m_2-1\} & \text{linke untere Ecke} \\ \{0, m_1-1; 0, L\} & \text{linker Rand} \end{cases}$$

identisch ist mit der Reaktion eines entsprechenden ortsinvarianten Systems auf das an Rändern mit einer geeigneten Extrapolationsvorschrift $f(u)$ erweiterte Eingangssignal.

$$u_0(k,l) = \begin{cases} u(k,l) & \text{für } (k,l) \in \{0,K;0,L\} \\ f(u) & \text{sonst} \end{cases} \quad (3.2-3)$$

Daraus folgt die signalunabhängig zu erfüllende Bedingungs-
gleichung

$$\begin{aligned} \tilde{y}(k,l) &= \sum_{\mu=-r_1}^{\alpha} \sum_{v=-r_2}^{\beta} \tilde{b}_{\mu v}(k,l) u(k-\mu, l-v) \\ &\stackrel{!}{=} \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} \cdot u_0(k-\mu, l-v) = y(k,l). \end{aligned} \quad (3.2-4)$$

Diese Vorschrift führt bei gegebenen Koeffizienten $\tilde{b}_{\mu v}(k,l)$ zu einer Extrapolationsvorschrift $f(u)$, wobei wir davon Gebrauch machen, daß im Innern des Bildes

$$\tilde{b}_{\mu v}(k,l) = b_{\mu v} \quad (3.2-5)$$

ist.

Umgekehrt erhält man die $\tilde{b}_{\mu v}(k,l)$ bei vorgegebenen $f(u)$ und $b_{\mu v}$.

Dadurch ist die Filterung eines endlichen Eingangsbildes mit ortsvarianten Randfiltern äquivalent mit der ortsinvarianten Filterung des am Rand erweiterten Eingangsbildes. Die Erweiterung, entsprechend der Extrapolationsvorschrift $f(u)$, beträgt

m_2 Punkte am oberen ,

r_2 Punkte am unteren,

r_1 Punkte am linken und

m_1 Punkte am rechten Rand.

Im Bild 3.3 ist dies schematisch dargestellt.

Eine Möglichkeit der Extrapolation wird zunächst am 1-D Fall erläutert.

Extrapoliert werde das 1-D Signal

$$u(k) = \sum_{\mu=0}^K u(\mu) \cdot \gamma_0(k-\mu) \quad (3.2-6)$$

$$k \in \{0, K\} \quad .$$

Im Bereich $k \in \{-K, 0\}$ und $k \in \{K, 2K\}$ werde das Signal derart fortgesetzt, daß gilt

$$u_0(k) = \begin{cases} 2u(0) - u(|k|) & -K \leq k \leq 0 \\ u(k) & 0 \leq k \leq K \\ 2u(K) - u(K-k) & K \leq k \leq 2K \end{cases} \quad (3.2-7)$$

Dies wird an einem Beispiel in Bild 3.4b verdeutlicht. Das damit erhaltene Eingangssignal $u_0(k)$ läßt sich zerlegen in

$$u_0(k) = u_r(k) + u_p(k) \quad .$$

Dabei ist $u_r(k)$ ein linear ansteigendes Signal mit dem Steigungsfaktor $\frac{u(K)-u(0)}{K}$ (Bild 3.4c), $u_p(k)$ kann als ein Ausschnitt aus einem periodischen Signal mit der Periode $2K$ (Bild 3.4d) gedeutet werden.

Sinngemäß läßt sich diese Extrapolationsvorschrift auf zweidimensionale Signale erweitern.

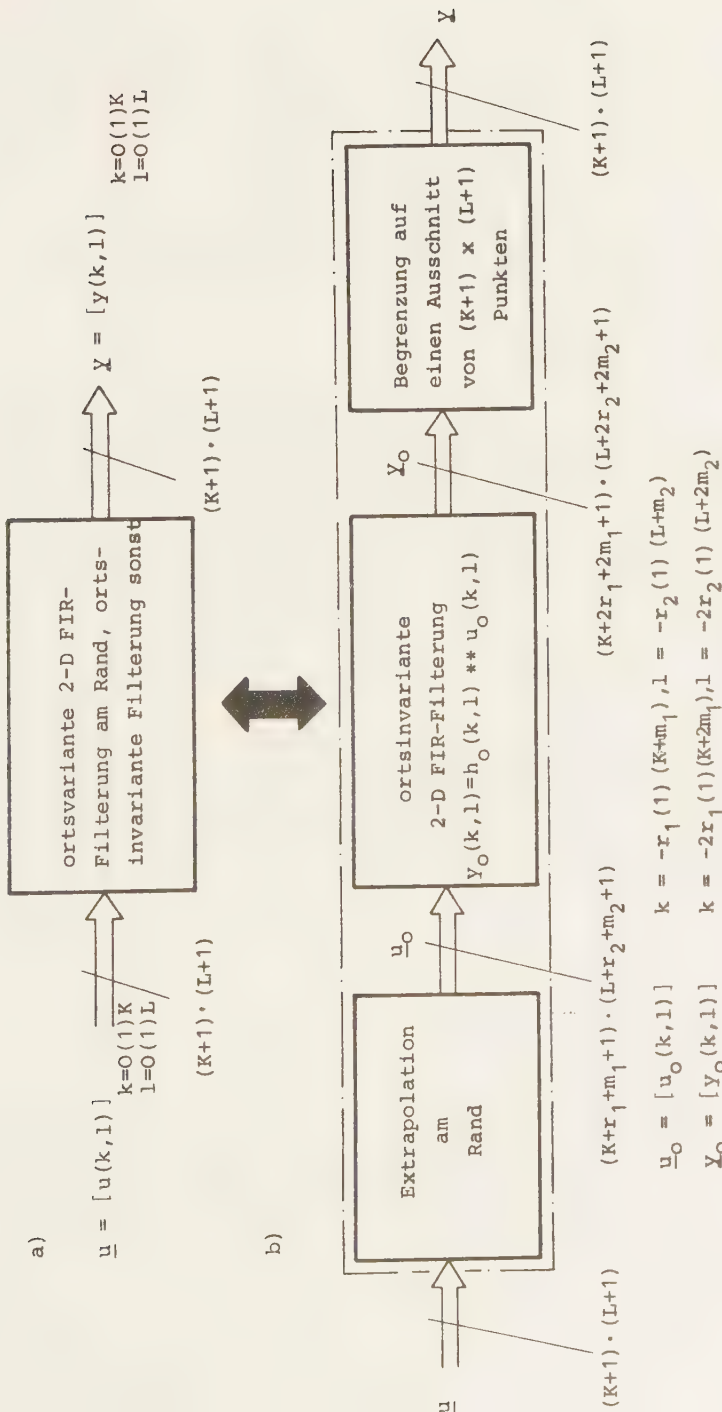


Bild 3.3: Zur Äquivalenz der ortsvarianten Randfilterung und der Filterung mit einem ortsinvarianten System

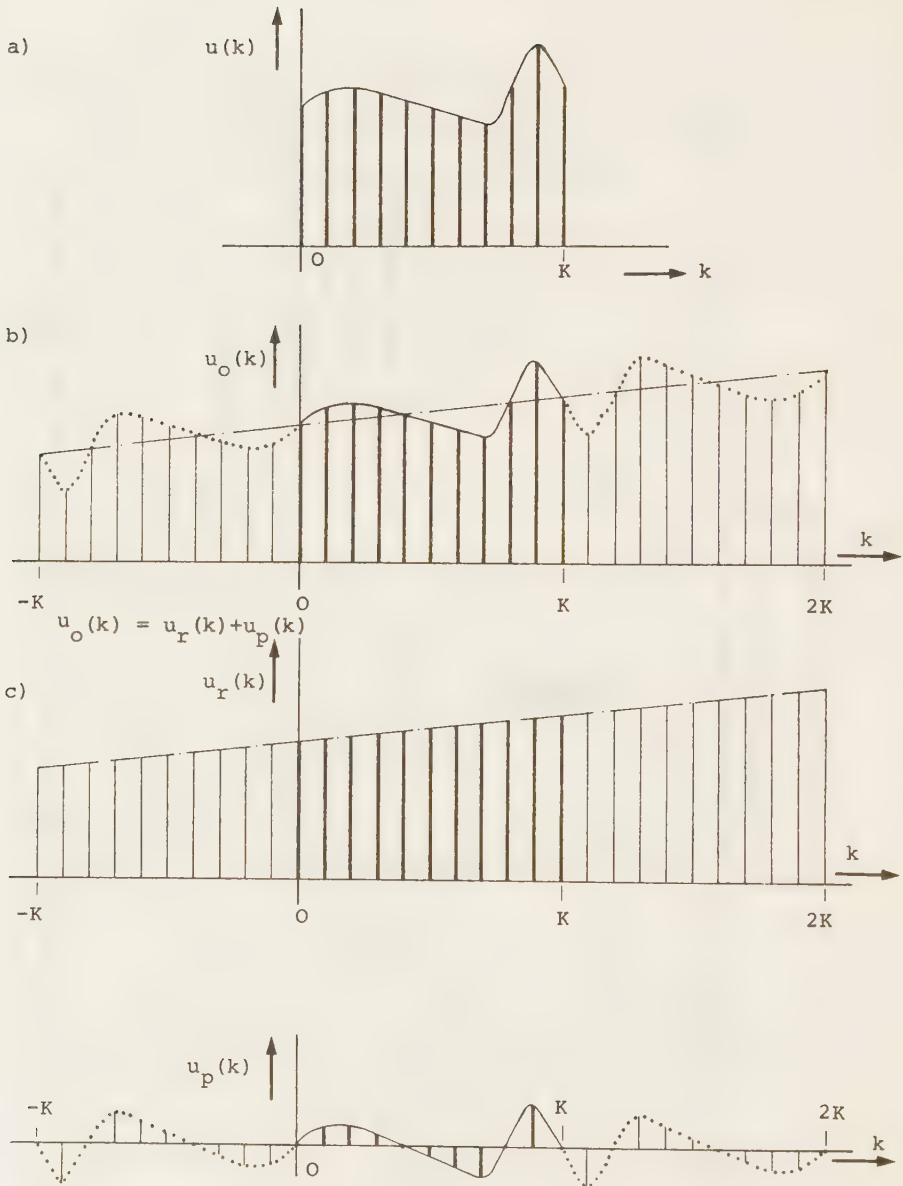


Bild 3.4: Zur Extrapolation eines 1-D-Signals

Für die in Bild 3.5 bezeichneten Bereiche ergibt sich

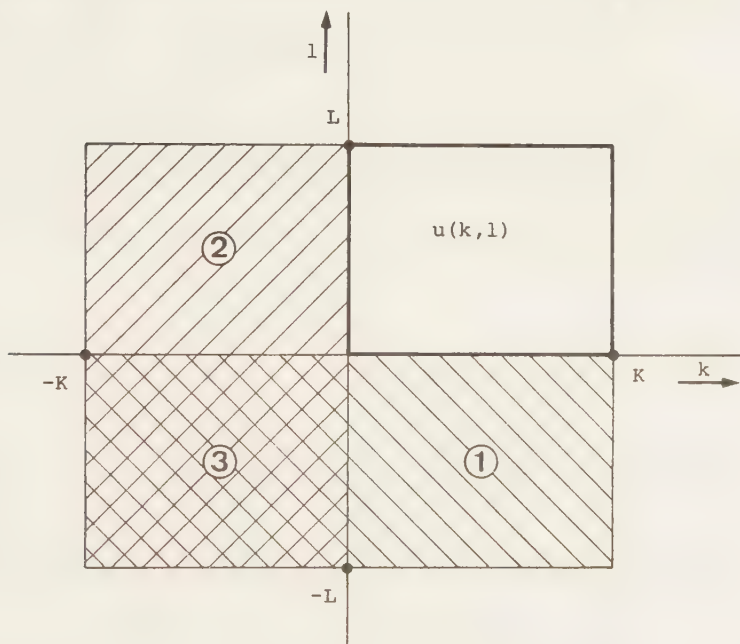


Bild 3.5: Unterschiedliche Extrapolationsbereiche
bei einem 2-D Eingangssignal

in Erweiterung von (3.2-7):

$$u_o(k, l) = \begin{cases} 2u(k, 0) - u(k, |l|) & 0 \leq k \leq K, -L \leq l \leq 0 \quad \textcircled{1} \\ 2u(0, l) - u(|k|, l) & -K \leq k \leq 0, 0 \leq l \leq L \quad \textcircled{2} \\ 2u(0, 0) - u(|k|, |l|) & -K \leq k \leq 0, -L \leq l \leq 0 \quad \textcircled{3} \end{cases} \quad (3.2-8)$$

Damit lassen sich die Koeffizienten der ortsvarianten Randfilter aus (3.2-4) bestimmen, wenn die Koeffizientenmatrix $\underline{B} = [b_{\mu\nu}]$ des ortsinvarianten Systems (3.1-4) gegeben ist. Erfüllt die Koeffizientenmatrix die Symmetriebedingungen

(Bild 3.6) $m_1=m_2=r_1=r_2=m$ und

$$b_{\mu\nu} = b_{|\mu||\nu|} \quad \mu, \nu \in \{-m, m\},$$

so ist die Impulsantwort dieses Filters oktandensymmetrisch, und es brauchen jeweils nur die Koeffizientenmatrizen der ortsvarianten Filter für einen Rand und eine Ecke berechnet werden. Alle anderen Filter erhält man durch entsprechende Drehung der Matrizen.

$$\underline{B} = \begin{bmatrix} \Delta & \blacksquare & \cdot & \blacksquare & \Delta \\ \blacksquare & x & + & x & \blacksquare \\ \cdot & + & o & + & \cdot \\ \blacksquare & x & + & x & \blacksquare \\ \Delta & \blacksquare & \cdot & \blacksquare & \Delta \end{bmatrix}$$

Bild 3.6: Koeffizientenmatrix eines oktandensymmetrischen FIR-Filters für den Fall $m_1=m_2=r_1=r_2=2$

Gleiche Koeffizienten sind durch das gleiche Symbol gekennzeichnet.

Als Beispiel sei der Laplaceoperator

$$\underline{B} = \begin{bmatrix} b_{-1,-1} & b_{-1,0} & b_{-1,1} \\ b_{0,-1} & b_{00} & b_{01} \\ b_{1,-1} & b_{1,0} & b_{11} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.2-9)$$

betrachtet.

Für die Koeffizientenmatrizen der ortsvarianten Randfilter ergibt sich mit (3.2-9) und (3.2-8) aus (3.2-4) für die Ecken

$$\underline{\underline{B}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.2-10a)$$

und für den linken Rand

$$\underline{\underline{B}} = \begin{bmatrix} 1 & 0 \\ -2 & 0 \\ 1 & 0 \end{bmatrix} \quad (3.2-10b)$$

In Bild 3.7 ist für ein Testsignal die Äquivalenz von ortsinvarianter Filterung mit extrapoliertem Eingangssignal und ortsvarianter Filterung verdeutlicht. Der Vergleich mit der ortsinvarianten Filterung ohne Erweiterung des Eingangsbildes zeigt, daß damit das Ein- und Ausschwingen am Rand (mit Ausnahme der Ecken) eliminiert wurde.

Für die Realisierung ist es unter Umständen einfacher, anstelle der ortsvarianten Filterung das extrapolierte Eingangssignal zu berechnen und anschließend ortsinvariant zu filtern.

3.3 Implementierung eines zweidimensionalen FIR-Filters als eindimensionales System

Wird die Filterung an einem Minicomputer durchgeführt, kann wegen des begrenzten Speicherplatzes im allgemeinen nicht das volle Ein- und Ausgangsbild im Hauptspeicher resident bleiben. Es wird deshalb sequentiell, im allgemeinen zeilenweise, aus einem Sekundärspeicher abgeholt und verarbeitet. Deshalb liegt es nahe, das zweidimensionale Filter für jeweils eine Zeile des Ein- und Ausgangsbildes zu definieren.

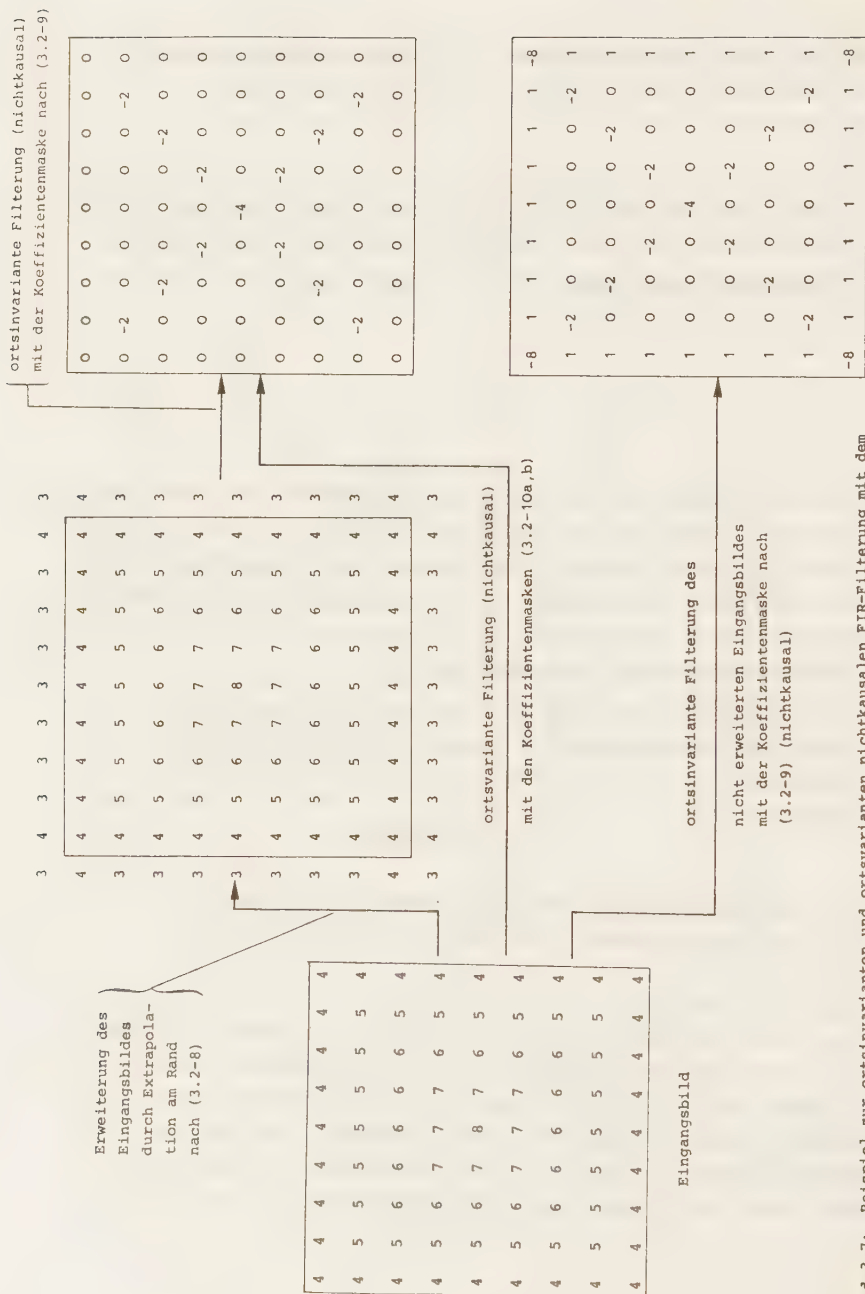


Bild 3.7: Beispiel zur ortsinvarianten und ortsvarianten nichtkausalen FIR-Filterung mit dem Laplace-Operator nach (3.2-9)

Ausgehend von (3.1-1)

$$y(k, l) = \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} \cdot u(k-\mu, l-v)$$

ergibt sich

$$y_l(k) = \sum_{v=-r_2}^{m_2} b_v(k) * u_{1-v}(k) \quad (3.3-1)$$

mit $l = -r_2, \dots, L+m_2$

$k = -r_1, \dots, K+m_1$.

Dabei bezeichnet $y_l(k)$ die 1-D-Folge der l -ten Zeile des Ausgangsbildes, $b_v(k)$ die v -te Spalte der Koeffizientenmatrix $\underline{B}$ nach (3.1-4) und $u_{1-v}(k)$ die $(1-v)$ -te Zeile des Eingangsbildes. Damit ist das zweidimensionale System auf ein 1-D-System mit mehreren Ein- und Ausgängen zurückgeführt [2]. Die Zahl der Ein- und Ausgänge ist nicht von vornherein festgelegt, sondern hängt von der Größe des zu berechnenden Bildausschnittes in k -Richtung ab. Durch eindimensionale $\mathcal{Z}$ -Transformation bzgl. k erhält man

$$\begin{aligned} Y_l(z_1) &= \sum_{v=-r_2}^{m_2} B_v(z_1) \cdot U_{1-v}(z_1) = \\ &= \sum_{v=0}^{m_2+r_2} B_{v-r_2}(z_1) \cdot U_{1+r_2-v}(z_1), \end{aligned} \quad (3.3-2)$$

$$\text{wobei } B_v(z_1) = \sum_{\mu=-r_1}^{m_1} b_{\mu v} \cdot z_1^{-\mu}$$

die Übertragungsfunktion des v -ten eindimensionalen nichtrekursiven Teilfilters bedeutet. Diese folgt auch unmittelbar aus der Übertragungsfunktion des zweidimensionalen FIR-Filters (3.1-2), da

$$H(z_1, z_2) = \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} z_1^{-\mu} z_2^{-v} = \sum_{v=-r_2}^{m_2} B_v(z_1) z_2^{-v} \quad (3.3-3)$$

ist.

Für die Realisierung eines allgemeinen 2-D nichtrekursiven Filters in direkter Struktur ergibt sich der in Bild 3.8 angegebene Signalflußgraph.

Aus $m_2 + r_2 + 1$ Zeilen des Eingangsbildes berechnet sich eine Zeile des Ausgangsbildes.

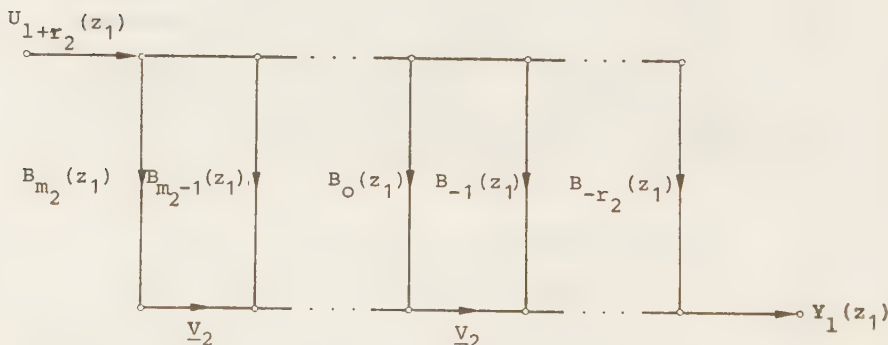


Bild 3.8: Signalflußgraph für ein allgemeines nichtrekursives Filter in direkter Form (1. kanonische Form). Das Verzögerungselement $\underline{V}_2$ speichert jeweils eine ganze Zeile.

3.4 Realisierung von 2-D FIR-Filtern durch schnelle Faltung [1,8]

Nach Gl. (2.3-8) ergibt sich die Ausgangsfolge $y(k,l)$ als nichtzyklische Faltung der 2-D Folge der Impulsantwort $h_o(k,l)$ mit der 2-D Eingangsfolge $u(k,l)$ in k und l Richtung. Dabei hat nach Bild 3.2 die Ausgangsfolge $y(k,l)$ die Dimension $(K+m_1+r_1+1) \times (L+m_2+r_2+1)$. Man kann die nichtzyklische Faltung durch eine zyklische Faltung ausführen, wenn man hinreichend viele Nullen ergänzt und die erweiterten Folgen als periodisch fortgesetzt betrachtet (Bild 3.9).

Dadurch erhält man zwei periodische Folgen

$$v_1(k,l), v_2(k,l), \quad k = (0,1,\dots) \bmod M_1$$

$$l = (0,1,\dots) \bmod M_2.$$

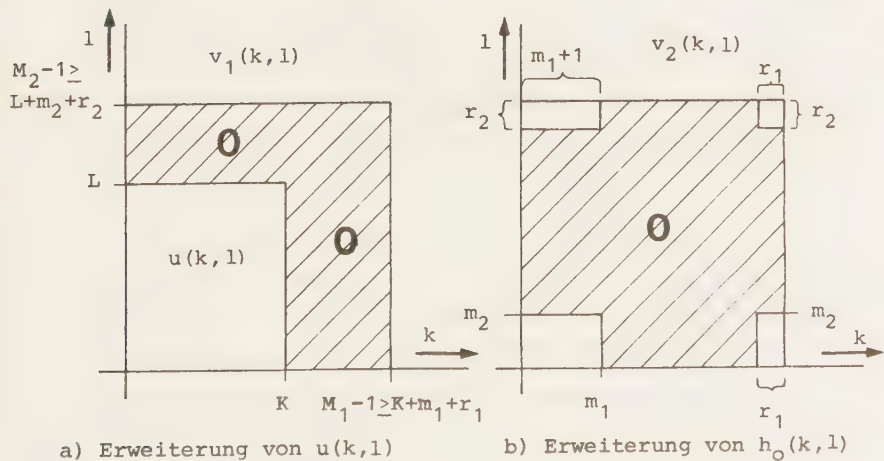


Bild 3.9: Realisierung einer nichtzyklischen Faltung als zyklische Faltung durch Erweiterung der Folgen $u(k,l)$ und $h_o(k,l)$

Bezeichnet $v_3(k, l)$ die periodisch fortgesetzte Ausgangsfolge von $y(k, l)$, so ergibt sich die zyklische Faltung

$$v_3(k, l) = v_1(k, l) \tilde{**} v_2(k, l) \quad (3.4-1)$$

$$k = (0, 1, \dots) \bmod M_1, \quad l = (0, 1, \dots) \bmod M_2.$$

Durch Anwendung einer geeigneten Transformation T lässt sich die zyklische Faltung nach Gl.(3.4-1) in eine punktweise Multiplikation der transformierten Folgen überführen. Durch Rücktransformation T^{-1} ergibt sich die gewünschte Ausgangsfolge

$$v_3(k, l) = T^{-1} \{ T\{v_1(k, l)\} \cdot T\{v_2(k, l)\} \}. \quad (3.4-2)$$

Die verwendeten Transformationen müssen eineindeutig sein.

Neben der am häufigsten verwendeten diskreten Fouriertransformation (DFT) finden auch andere Transformationen wie z.B. Haar, Slant, Hadamard [5,8] oder zahlentheoretische Transformationen [18,19] Anwendung. Lassen sich diese Transformationen nach einem schnellen Transformationsalgorithmus berechnen (z.B. FFT-Algorithmus [1]), so bezeichnet man diese Berechnungsweise als **s c h n e l l e F a l t u n g**.

Führt man die zyklische Erweiterung der Eingangsfolge $u(k, l)$ in $v_1(k, l)$ nicht durch, so entstehen bei Ausführung der schnellen Faltung Überlappungsfehler [8]. Dies ist in Bild 3.10 dargestellt.

Die innerhalb des nichtschraffierten Bereichs von $v_3(k, l)$ liegenden Punkte sind richtig und entsprechen den Werten der Ausgangsfolge $y(k, l) = u(k, l) ** h_0(k, l)$.

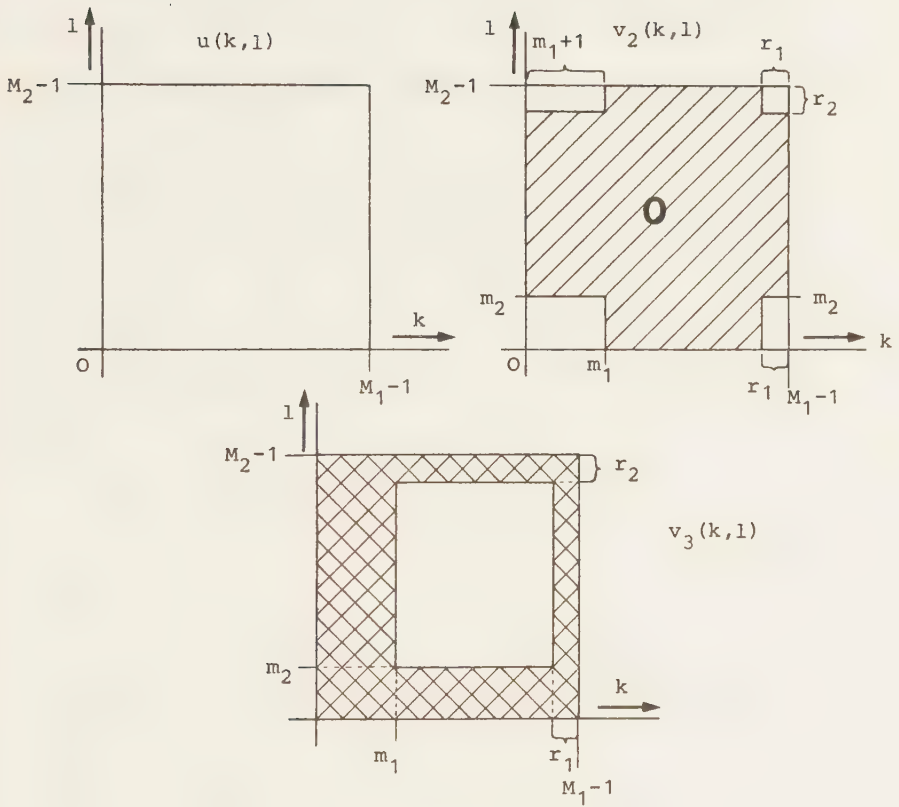


Bild 3.10: Zum Überlappungsfehler bei der 2-D zyklischen Faltung

Im nächsten Abschnitt wird die 2-D FFT und deren Implementierung an einem Minicomputer genauer behandelt.

3.5 Zweidimensionale FFT, Implementierung an einem Minicomputer

Die diskrete Fouriertransformierte (DFT) eines zweidimensionalen periodischen Signals $v(k,l)$ $k=(0,1,\dots) \bmod M_1$ $l=(0,1,\dots) \bmod M_2$ wird beschrieben durch [3]

$$\begin{aligned} \text{DFT}\{v(k,l)\} &= V(\kappa,\mu) = \sum_{k=0}^{M_1-1} \sum_{l=0}^{M_2-1} v(k,l) w_{M_1}^{\kappa k} \cdot w_{M_2}^{\mu l} \\ &= \sum_{l=0}^{M_2-1} \left[\sum_{k=0}^{M_1-1} v(k,l) w_{M_1}^{\kappa k} \right] w_{M_2}^{\mu l} \quad (3.5-1) \\ &\quad -j \frac{2\pi}{M_1} \quad -j \frac{2\pi}{M_2} \end{aligned}$$

wobei $w_{M_1} = e^{-j \frac{2\pi}{M_1}}$, $w_{M_2} = e^{-j \frac{2\pi}{M_2}}$,

$k, \kappa = 0, 1, \dots, M_1-1$; $l, \mu = 0, 1, \dots, M_2-1$ ist.

Ein Vergleich mit der Z-Transformierten von $v(k,l)$ zeigt, daß

$$\text{DFT}\{v(k,l)\} = Z\{v(k,l)\} \Big|_{z_1 = e^{j \frac{2\pi \kappa}{M_1}}, z_2 = e^{j \frac{2\pi \mu}{M_2}}}$$

gilt.

Für die Umkehrtransformation ergibt sich

$$\text{DFT}^{-1}\{V(\kappa,\mu)\} = v(k,l) = \frac{1}{M_1 M_2} \cdot \sum_{\kappa=0}^{M_1-1} \sum_{\mu=0}^{M_2-1} V(\kappa,\mu) w_{M_1}^{-\kappa k} w_{M_2}^{-\mu l} \quad (3.5-2)$$

Durch die Schreibweise von (3.5-1) wird die Möglichkeit deutlich, die zweidimensionale Transformation in eindimensionale Transformationen zu zerlegen. In Abschnitt 3.5.2 wird dies ausführlich behandelt.

Wir merken zunächst an, daß bei den hier gegebenen reellen Eingangsfolgen mit geradzahligem M_1 oder M_2 durch Ausnutzung der Symmetrieeigenschaften der DFT [1,8] die Transformationslänge in einer Richtung auf die Hälfte verkürzt werden kann. Dazu faßt man je zwei aufeinanderfolgende Werte zu einem komplexen Wert zusammen. Ist M_1 geradzahlig, lassen sich je zwei aufeinanderfolgende Werte einer Zeile zusammenfassen.

Aus der reellen 2-D Folge $v(k,l)$ wird die komplexwertige Folge

$$v_c(m,l) = v(2m,l) + j v(2m+1,l) \quad (3.5-3)$$

$$m=0,1,\dots, \frac{M_1}{2}-1, \quad l=0,1,\dots, M_2-1$$

gebildet. Mit

$$\text{DFT}\{v_c(m,l)\} = V_c(\kappa,\mu) = \sum_{m=0}^{\frac{M_1}{2}-1} \sum_{l=0}^{M_2-1} v_c(m,l) w_{\frac{M_1}{2}}^{\kappa m} \cdot w_{M_2}^{\mu l} \quad (3.5-4)$$

$$\kappa=0,1,\dots, \frac{M_1}{2}-1, \quad \mu=0,1,\dots, M_2-1$$

ergibt sich die gewünschte 2-D DFT der Folge $v(k,l)$ zu

$$\begin{aligned} V(\lambda,\mu) &= \frac{1}{2} [V_c(\lambda,\mu) + V_c^*\left(\frac{M_1}{2}-\lambda, M_2-\mu\right)] \\ &\quad - \frac{1}{2} j w_{M_1}^{\lambda} [V_c(\lambda,\mu) - V_c^*\left(\frac{M_1}{2}-\lambda, M_2-\mu\right)]. \end{aligned} \quad (3.5-5)$$

Wegen der Periodizität ist $V_c(\frac{M_1}{2}, \mu) = V_c(0, \mu)$.

Damit gilt (3.5-5) für $\lambda=0,1,\dots,\frac{M_1}{2}$, $\mu=0,1,\dots,M_2-1$.
Die restlichen Werte erhält man aus der Beziehung

$$V(\lambda, \mu) = V^*(M_1 - \lambda, M_2 - \mu). \quad (3.5-6)$$

Alternativ dazu lassen sich stattdessen auch zwei reellwertige Zeilen zu einer komplexwertigen Zeile der Eingangssignalmatrix $\underline{V} = [v(k,l)]$ zusammenfassen.

Zur Reduktion des notwendigen Speicherplatzes und der Minimierung der Daten-Ein/Ausgabezeit ist es für die Implementierung an einem Minicomputer notwendig, diese Zusammenfassung stets vorzunehmen. Die zusätzlichen Multiplikationen und Additionen für die Umrechnung nach (3.5-5) spielen dabei eine untergeordnete Rolle.

Das Hauptproblem bei der Durchführung der Transformation liegt darin, daß für $M_1 \cdot M_2 > 2^{12}$ normalerweise der Hauptspeicher oder aber der Adreßraum (bei 16 bit Maschinen ohne virtuelle Adressierung) nicht ausreicht. Zudem ist das Eingangssignal sequentiell (i. allg. zeilenweise) auf einem Massenspeicher (z.B. Magnetplatte, Magnetband) abgespeichert. Im Gegensatz zum 1-D Fall spielt deshalb neben der Zeit für die Berechnung der Transformation, welche im wesentlichen durch die Zahl der Multiplikationen und Additionen bestimmt wird, die Zeit für den Datenaustausch zwischen Haupt- und Massenspeicher eine entscheidende Rolle.

Die Datentransferzeit kann dabei je nach der Zahl der mit einem Sekundärspeicherzugriff in den Hauptspeicher eingelesenen Zeilen (oder Spalten) und der Transformationslänge über 80% der Gesamtrechnzeit betragen. In Bild 3.11 ist dies für die Zeilen/Spalten-FFT, die in Abschnitt 3.5.2 näher untersucht wird, graphisch dargestellt. N_B gibt dabei als Maß für den im Hauptspeicher verfügbaren Hauptspeicherplatz die Anzahl der gespeicherten reellen Zahlen an.

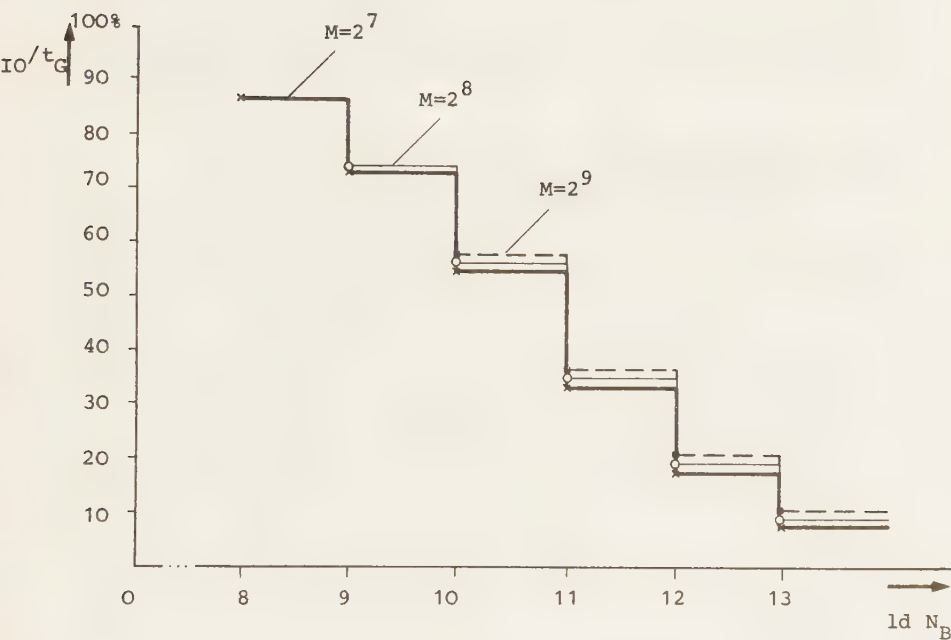


Bild 3.11: Prozentualer Anteil der Datentransferzeit t_{IO} an der Gesamtrechenzeit t_G in Abhängigkeit der Transformationslänge $M_1=M_2=M$ und der Blocklänge N_B für die Zeilen/Spalten-FFT nach Kap.3.5.2

Für die Beurteilung der Leistungsfähigkeit eines Algorithmus zur 2-D-FFT mit einem Minicomputer ist daher neben der Zahl der arithmetischen Operationen die Zahl der Zugriffe auf den Sekundärspeicher eine wesentliche Größe. Die Effizienz des jeweiligen Verfahrens ist dabei wesentlich von der Konfiguration des verwendeten Minicomputers abhängig.

3.5.1 Maschinenabhängige Größen zur Beurteilung der Leistungsfähigkeit von Algorithmen

Bestimmende Größen für die Beurteilung der Leistungsfähigkeit von Algorithmen zur Bildverarbeitung sind:

1. Die Größe des verfügbaren Datenspeicherplatzes im Hauptspeicher. Das entspricht der Zahl der gleichzeitig speicherbaren Zeilen (oder Spalten) der Signalmatrix. Im weiteren beschreibt N_B die Zahl der gleichzeitig speicherbaren reellen Datenworte.
2. Die typischen Zeiten für den Sekundärspeicher, in dem die Ein- und Ausgangsdaten gespeichert werden.

Diese sind:

T_{acc} = mittlere Zugriffszeit auf einen Datenblock

T_{tran} = Transferzeit für eine reelle Zahl von Sekundär- zum Hauptspeicher oder umgekehrt.

3. Die typischen Zeiten für die Durchführung von arithmetischen Verknüpfungen in der arithmetischen Einheit (bei der PDP 11/45: Gleitkommaprozessor). Einbezogen wird dabei im allgemeinen die Zeit für den Datentransfer zwischen Hauptspeicher und arithmetischer Einheit.
- Diese sind:

T_{mult} = Zeit für eine reelle Multiplikation

T_{add} = Zeit für eine reelle Addition.

Für die PDP 11/45 wurden in Assembler bei Verwendung von RK05-Platten als Sekundärspeicher folgende Zeiten ermittelt:

$$T_{\text{acc}} = 60\,000 \text{ } \mu\text{sec}$$

$$T_{\text{tran}} = 24 \text{ } \mu\text{sec}$$

$$T_{\text{add}} = 24 \text{ } \mu\text{sec}$$

$$T_{\text{mult}} = 27 \text{ } \mu\text{sec.}$$

3.5.2 Zerlegung der 2-D-DFT in 1-D-DFT's

Die 2-D-DFT nach Gl.(3.5-4) läßt sich auf die Berechnung der DFT von eindimensionalen Folgen zurückführen.

Mit

$$\tilde{V}_C(\kappa, l) = \text{DFT}\{V_C(m, l)\} = \sum_{m=0}^{M_1-1} V_C(m, l) w_{M_1}^{\kappa m} \quad (3.5-7)$$

$$\kappa = 0, 1, \dots, \frac{M_1}{2} - 1, \quad l = 0, 1, \dots, M_2 - 1$$

ergibt sich aus (3.5-4)

$$V_C(\kappa, \mu) = \text{DFT}\{\tilde{V}_C(\kappa, l)\} = \sum_{l=0}^{M_2-1} \tilde{V}_C(\kappa, l) w_{M_2}^{\mu l} \quad (3.5-8)$$

$$\kappa = 0, 1, \dots, \frac{M_1}{2} - 1, \quad \mu = 0, 1, \dots, M_2 - 1.$$

Gl.(3.5-7) beschreibt eine eindimensionale DFT der Länge $\frac{M_1}{2}$ längs der l-ten Zeile und (3.5-8) eine eindimensionale DFT der Länge M_2 längs der κ -ten Spalte. Die DFT's können nach einem der bekannten FFT-Algorithmen (Cooley-Tuckey, Singleton, Winograd) [20] berechnet werden.

Nach der zeilenweisen Berechnung der eindimensionalen DFT der Länge $\frac{M_1}{2}$ ist es zweckmäßig, bereits die Umrechnung auf die eindimensionale DFT der Länge M_1 für reelle Werte vorzunehmen.

Es gilt:

$$\begin{aligned} \tilde{V}(\kappa, 1) = & \frac{1}{2} \left\{ [\tilde{V}_C(\kappa, 1) + \tilde{V}_C^* \left(\frac{M_1}{2} - \kappa, 1 \right)] \right. \\ & \left. - j w_{M_1}^{\kappa} [\tilde{V}_C(\kappa, 1) - \tilde{V}_C^* \left(\frac{M_1}{2} - \kappa, 1 \right)] \right\}, \\ \kappa = & 0, 1, \dots, \frac{M_1}{2}, \quad 1 = 0, 1, \dots, M_2 - 1. \end{aligned} \quad (3.5-9)$$

Da wegen

$$\begin{aligned} \tilde{V}(0, 1) &= \operatorname{Re} \{ \tilde{V}_C(0, 1) \} + \operatorname{Im} \{ \tilde{V}_C(0, 1) \} \\ \tilde{V}\left(\frac{M_1}{2}, 1\right) &= \operatorname{Re} \{ \tilde{V}_C(0, 1) \} - \operatorname{Im} \{ \tilde{V}_C(0, 1) \} \end{aligned} \quad (3.5-10)$$

$\tilde{V}(0, 1)$ und $\tilde{V}\left(\frac{M_1}{2}, 1\right)$ stets reellwertig sind, können diese beiden Werte zu einem komplexen Wert zusammengefaßt werden. Es ergibt sich die 2-D-Folge

$$\tilde{X}(\kappa, 1) = \begin{cases} \tilde{V}(0, 1) + j \tilde{V}\left(\frac{M_1}{2}, 1\right) & \kappa=0 \\ \tilde{V}(\kappa, 1) & \kappa=1, 2, \dots, \frac{M_1}{2} - 1 \end{cases} \quad (3.5-11)$$

Vor der Ausführung der spaltenweisen 1-D DFT der Länge M_2 muß die Matrix $\tilde{\underline{X}} = [\tilde{X}(\kappa, 1)]$

$\kappa = 0, 1, \dots, \frac{M_1}{2} - 1$ transponiert werden. In der Literatur wurden Verfahren zur Matrixtransponierung eingehend untersucht (z.B. [21-23]). Am bekanntesten ist das von Eklundh [21] entwickelte Verfahren für quadratische $M \times M$ -Matrizen mit M als Zweierpotenz. Dies ist jedoch wegen der verkürz-

ten Transformationslänge bei reellen Eingangsdaten im allgemeinen nicht anwendbar. Zudem ist es effektiver, das notwendige 'bitreversal' mit in den Transponierungsalgorithmus einzubeziehen. Für eine Radix-2-FFT wird hier ein neues Verfahren angewandt.

Das Verfahren erfordert

$$n = \lg \frac{2^{M_1 M_2}}{N_B} \quad (3.5-12)$$

Schritte mit

$$m = \frac{4^{M_1 M_2}}{N_B} - \begin{cases} 1 & \text{beim 1. und letzten Schritt} \\ 2 & \text{sonst} \end{cases} \quad (3.5-13)$$

Ein- oder Ausleseoperationen eines Datensatzes bestehend aus $N_B/2$ reellen Werten.

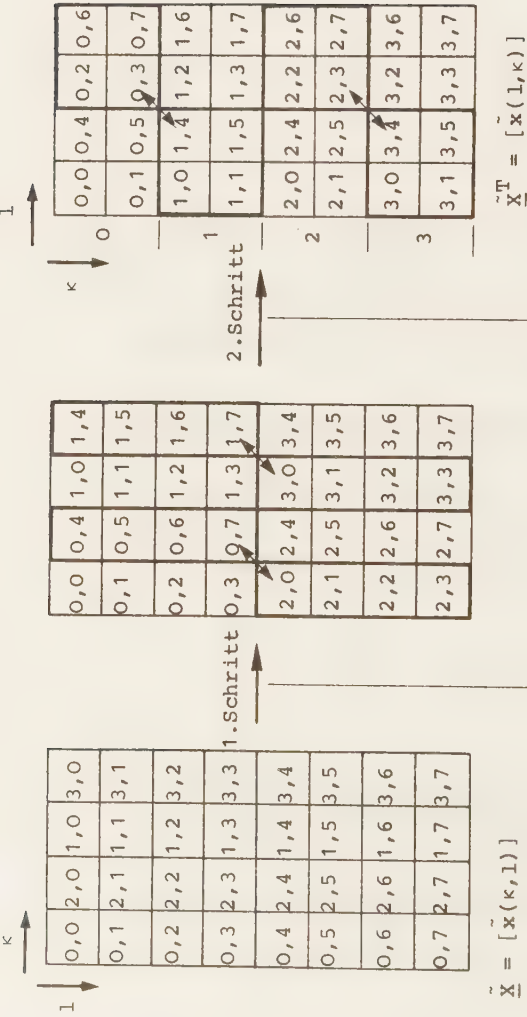
Voraussetzung dabei ist, daß M_1, M_2, N_B Zweierpotenzen sind und $N_B \geq 2 \max\{M_1, M_2\}$ gilt.

Am Beispiel $M_1=M_2=8$, $N_B=32$ ist das Verfahren in Bild 3.12 erläutert. Nach der Transponierung stehen die neuen Zeilen in 'bitreversed'-Reihenfolge, wodurch sich nach der eindimensionalen DFT

$$X(\mu, \kappa) = \text{DFT}\{\tilde{X}(l, \kappa)\} = \sum_{l=0}^{M_2-1} \tilde{X}(l, \kappa) w_{M_2}^{l\mu} \quad (3.5-14)$$

mit Bitreversal am Eingang die richtige Reihenfolge am Ausgang ergibt. Die Ausgangsmatrix bleibt transponiert, was meistens nicht weiter stört.

Aus $X(\mu, 0)$ berechnet sich die 0-te und $\frac{M_1}{2}$ Zeile der DFT der 2-D-Folge zu



7 IO-Operationen

7 IO-Operationen

$$m = 2 \cdot \left(\frac{M_1 M_2}{N_B} - 1 \right) = 14$$

Zahl der Ein/Ausgabeoperationen:

$$n = \text{ld} \frac{M_1 M_2}{N_B} = 2$$

Zahl der Schritte

Bild 3.12: Beispiel zur Matrixtransponierung mit "bitreversal" für $M_1=M_2=8$, $N_B=32$

$$V(0, \mu) = \frac{1}{2} [\operatorname{Re}\{X(\mu, 0)\} + \operatorname{Re}\{X(M_2 - \mu, 0)\} \\ + j (\operatorname{Im}\{X(\mu, 0)\} - \operatorname{Im}\{X(M_2 - \mu, 0)\})] \quad (3.5-15a)$$

$$V\left(\frac{M_1}{2}, \mu\right) = \frac{1}{2} [\operatorname{Im}\{X(\mu, 0)\} + \operatorname{Im}\{X(M_2 - \mu, 0)\} \\ - j (\operatorname{Re}\{X(\mu, 0)\} - \operatorname{Re}\{X(M_2 - \mu, 0)\})] \quad (3.5-15b)$$

und für $\kappa = 1, \dots, \frac{M_1}{2} - 1$ gilt

$$V(\kappa, \mu) = X(\mu, \kappa) \quad \mu = 0, 1, \dots, M_2 - 1. \quad (3.5-15c)$$

Die restlichen Werte ergeben sich nach (3.5-6).

Das hier vorgestellte Verfahren wurde in einer Studienarbeit [24] an der PDP 11/45 programmiert.

3.5.3 Gesamtrechenaufwand für die 2-D-FFT

Nach dem beschriebenen Verfahren mit Matrixtransponierung ergibt sich zur Berechnung von (3.5-1) unter der Voraussetzung, daß M_1, M_2, N_B Zweierpotenzen sind und N_B reelle Werte im Hauptspeicher Platz finden für $N_B \geq 2 \cdot \max\{M_1, M_2\}$ folgender Gesamtrechenaufwand:

- Zahl der notwendigen Datentransfers vom und zum Sekundärspeicher mit einer Blocklänge von $\frac{N_B}{2}$ reellen Werten:

$$N_{IO} = 2 \cdot \lg \left(\frac{2M_1 M_2}{N_B} \right) \left(\frac{2M_1 M_2}{N_B} - 1 \right) + 3 \quad (3.5-16)$$

- Zahl der komplexen Multiplikationen

$$N_{\text{mult}} = \frac{M_1 M_2}{4} \cdot \text{ld}(2 \cdot M_1 M_2) \quad (3.5-17)$$

- Zahl der komplexen Additionen

$$N_{\text{add}} = \frac{M_1 M_2}{2} \cdot \text{ld}(4 \cdot M_1 M_2) \quad (3.5-18)$$

Dabei sind die erforderlichen Multiplikationen und Additionen zur Spektrumumrechnung nach (3.5-15a-c) mit enthalten.

Mit den in 3.5.1 definierten maschinenabhängigen Bestimmungsgrößen ergibt sich als Dateneinausgabezeit t_{IO} :

$$t_{\text{IO}} = \left[2 \cdot \text{ld} \left(\frac{2 \cdot M_1 M_2}{N_B} \right) \left(\frac{2 \cdot M_1 M_2}{N_B} - 1 \right) + 3 \right] \left(T_{\text{acc}} + \frac{N_B}{2} T_{\text{tran}} \right) \quad (3.5-19)$$

Unter der Voraussetzung, daß 1 komplexe Multiplikation 4 reellen Multiplikationen und 2 reellen Additionen und eine komplexe Addition 2 reellen Additionen entspricht, ergibt sich mit (3.5-17) und (3.5-18) die Zeit für die arithmetische Berechnung t_A

$$t_A = \frac{M_1 M_2}{2} \left[2 \text{ld}(2 M_1 M_2) T_{\text{mult}} + \text{ld}(32 \cdot M_1^3 M_2^3) \cdot T_{\text{add}} \right] \quad (3.5-20)$$

In Bild 3.11 ist für $M_1 = M_2 = M$ das prozentuale Verhältnis der Dateneinausgabezeit zur Gesamtrechnenzeit

$$t_G = t_{\text{IO}} + t_A$$

in Abhängigkeit von N_B dargestellt.

Eine Reduktion der Zeit für die arithmetische Berechnung der DFT kann erreicht werden, wenn statt der zeilen- und spaltenweisen FFT mit Matrixtransponierung die 'Vektor-Radix'-FFT [25] angewandt wird. Allerdings benötigt dieses Verfahren mehr Datentransfers, nämlich [25]

$$N_{IO} = 2 \cdot \lg M_1 \cdot \left(\frac{2M_1 M_2}{N_B} - 1 \right).$$

Deshalb bringt die Anwendung dieses Verfahrens nur dann Vorteile, wenn im Hauptspeicher jeweils nur 2 Zeilen abgespeichert werden können ($N_B = 2M_1$), oder aber der Hauptspeicherplatz ausreicht, um die zweidimensionale DFT vollständig auszuführen [24].

3.6 Aufwandsvergleich für die Realisierung nichtrekursiver Filter an einem Minicomputer

In diesem Kapitel soll der Realisierungsaufwand für die Implementierung der nichtrekursiven Filterung mit Hilfe der DFT und der direkten Realisierung (Bild 3.1) verglichen werden. Dazu werden folgende Voraussetzungen gemacht:

- Es sollen $N \times N$ Werte einer zweidimensionalen reellen Folge mit einem nichtrekursiven Filter (3.1-1) mit $m_1 = m_2 = r_1 = r_2 = m$, m ganzzahlig, gefiltert werden.
- Das Filter soll Nullphase besitzen, für die Impulsantwort soll also gelten

$$h_0(\mu, \nu) = h_0(-\mu, -\nu) \quad . \quad (3.6-1)$$

- Vom Ausgangsbild sollen $N \times N$ Werte berechnet werden. Ferner sei zum Vergleich mit der zweidimensionalen FFT der dafür günstigste Fall angenommen, daß

$M = N+2m$ eine Zweierpotenz ist.

- Im Hauptspeicher sollen N_B reelle Werte gespeichert werden können, wobei N_B ebenfalls eine Zweierpotenz ist und

$$N_B > 2(m+1)N \quad \text{sein soll.}$$

- Es wird ferner angenommen, daß das Eingangssignal zeilenweise auf dem Sekundärspeicher abgespeichert ist.

Unter diesen Voraussetzungen ergibt sich für die direkte Filterung nach Bild 3.1 folgender Rechenaufwand pro Ausgangswert:

$$N_{\text{mult}} = (m+1)+m(2m+1) \text{ reelle Multiplikationen,}$$

$$N_{\text{add}} = (2m+1)^2 - 1 = 4m(m+1) \text{ reelle Additionen.}$$

Deshalb beträgt die Zeit für die arithmetische Berechnung

$$t_A = N^2 \{ N_{\text{mult}} \cdot T_{\text{mult}} + N_{\text{add}} \cdot T_{\text{add}} \}$$

$$= N^2 \{ [(m+1)+m(2m+1)] \cdot T_{\text{mult}} + 4m(m+1) \cdot T_{\text{add}} \}. \quad (3.6-2)$$

Unter der Annahme, daß $r = \left\lceil \frac{N_B}{N} \right\rceil$ Zeilen ($[x]$ ist die größte ganze Zahl, kleiner oder gleich x) im Hauptspeicher speicherbar sind, werden jeweils $r-(2m+1)$ Zeilen vom Sekundärspeicher eingelesen, verarbeitet und anschließend wieder abgespeichert.

Damit ergeben sich

$$N_{IO} = 2 \cdot \left\langle \frac{N}{r - (2m+1)} \right\rangle \approx 2 \cdot \left\langle \frac{N^2}{N_B - (2m+1)N} \right\rangle$$

Ein- Ausgabeoperationen.

($\langle x \rangle$ ist die kleinste ganze Zahl, größer oder gleich x).
Da insgesamt N^2 Werte ein- und ausgelesen werden müssen, ergibt sich als Zeit für die Datenein- und ausgabe

$$t_{IO} = 2 \cdot \left\langle \frac{N^2}{N_B - (2m+1)N} \right\rangle T_{acc} + 2 \cdot N^2 \cdot T_{tran} . \quad (3.6-3)$$

Für die Realisierung durch Schnelle Faltung gelten nachfolgende Betrachtungen:

Unter der Voraussetzung, daß die DFT der Impulsantwort bereits vorberechnet auf dem Sekundärspeicher abgelegt ist und die damit notwendigen Einleseoperationen gerade dadurch kompensiert werden, daß bei der anschließenden Rücktransformation keine Einleseoperationen erforderlich sind, ergibt sich der zweifache Aufwand einer DFT der reellen zweidimensionalen Folge von der Dimension $N \times N$ und zusätzlich $M(\frac{M}{2}+1)$ komplexe Multiplikationen. Mit $M_1=M_2=M$ erhält man aus (3.5-19) als Datentransferzeit

$$t_{IO} = \left[4 \cdot \lg \left(\frac{2 \cdot M^2}{N_B} \right) \left(\frac{2 \cdot M^2}{N_B} - 1 \right) + 6 \right] (T_{acc} + \frac{N_B}{2} T_{tran}) \quad (3.6-4)$$

und als Zeit für die arithmetische Berechnung mit (3.5-20)

$$t_A = (4 \cdot M^2 \cdot \lg 2M + 4M) T_{mult} + (6 \cdot M^2 \cdot \lg 2M + 2M) T_{add} . \quad (3.6-5)$$

In der Bildverarbeitung werden sehr häufig nichtrekursive Filter mit quadratensymmetrischer Impulsantwort

$$h_o(\mu, \nu) = h_o(|\mu|, |\nu|) \quad (3.6-6)$$

verwendet. In diesem Fall reduziert sich die Zahl der arithmetischen Operationen bei Anwendung der eindimensionalen Realisierung nach 3.3 auf

$N_{\text{mult}} = (m+1)^2$ reelle Multiplikationen pro Wert
und

$N_{\text{add}} = 2m(m+2)$ reelle Additionen pro Wert.

In Bild 3.13 ist für beide Fälle nach Gl.(3.6-1) und Gl.(3.6-6) die Gesamtrechnenzeit unter Verwendung der in 3.5.1 angegebenen Zeitkonstanten für die PDP 11/45 in Abhängigkeit vom Filtergrad $n \times n$, $n=2m$ für $N_B = 2^{13}$ aufgetragen. Aus dem Diagramm läßt sich erkennen, daß für $M=2^8$ die schnelle Faltung bei nichtrekursiven Filtern, die Gl.(3.6-1) erfüllen ab einem Filtergrad von 8×8 , bei Filtern, die zusätzlich quadrantensymmetrische Impulsantwort besitzen (3.6-6), ab einem Filtergrad von 12×12 effektiver als die direkte Realisierung ist.

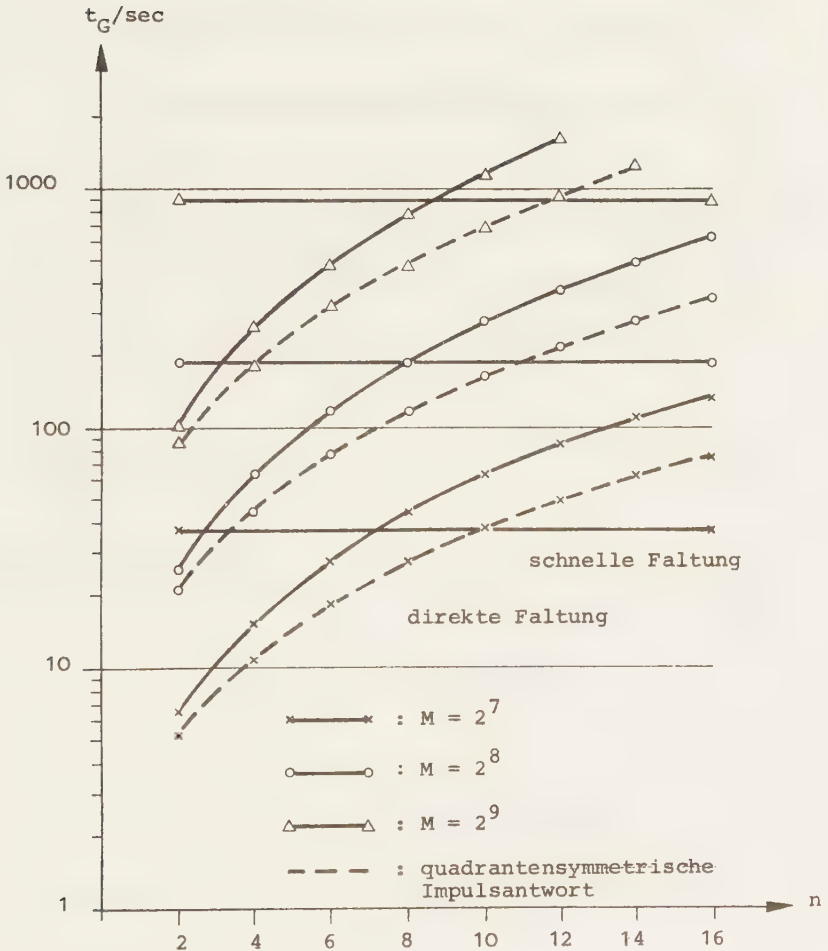


Bild 3.13: Gesamtrechnenzeit t_G für die schnelle und direkte Faltung in Abhängigkeit vom Filtergrad n und der FFT-Länge M für $N_B = 2^{13}$

4. Zweidimensionale rekursive digitale Filter

4.1 _Grundlegende Definitionen und Eigenschaften_

4.1.1 Differenzengleichung, Übertragungsfunktion

Lineare ortsinvariante rekursive digitale Filter haben im Gegensatz zu nichtrekursiven Filtern eine theoretisch unendlich ausgedehnte Impulsantwort. Sie werden deshalb auch als IIR-Filter (infinite impulse response) bezeichnet. Derartige Filter lassen sich stets durch eine partielle Differenzengleichung nach (2.3-15)

$$\sum_{\mu=-q_1}^{n_1} \sum_{v=-q_2}^{n_2} c_{\mu v} y(k-\mu, l-v) = \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} u(k-\mu, l-v) \quad (4.1-1)$$

beschreiben und besitzen demnach eine rationale Übertragungsfunktion

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)} = \frac{\sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} z_1^{-\mu} z_2^{-v}}{\sum_{\mu=-q_1}^{n_1} \sum_{v=-q_2}^{n_2} c_{\mu v} z_1^{-\mu} z_2^{-v}}, \quad (4.1-2)$$

wobei $c_{00}=1$ sein soll.

Ein derart allgemeines nichtkausales Filter ist nicht rekursiv berechenbar und kann nur in der in Kap. 4.3.1 angegebenen Weise realisiert werden.

Rekursiv berechenbare Filter sind

- V i e r t e l e b e n e n f i l t e r mit kausalem Nennerpolynom, d.h. $q_1=q_2=0$
- H a l b e b e n e n f i l t e r mit semikausalem Nennerpolynom, d.h. $q_2=0$ und $c_{\mu 0}=0$ für $\mu < 0$

für Filter in ++ Rekursionsrichtung bei zeilenweiser Verarbeitungsreihenfolge (siehe 4.1.2).

Das Zählerpolynom hingegen braucht nicht kausal zu sein.

Damit ergeben sich folgende Differenzengleichungen und Übertragungsfunktionen für $c_{00}=1$.

Viertelebenenfilter mit i.allg. nichtkausalem Zählerpolynom

$$y(k,l) = \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} u(k-\mu, l-v) - \sum_{\mu=0}^{n_1} \sum_{v=0}^{n_2} c_{\mu v} y(k-\mu, l-v) \quad (4.1-3)$$

($\mu+v \neq 0$)

$$H(z_1, z_2) = \frac{\sum_{\mu=r_1}^{m_1} \sum_{v=r_2}^{m_2} b_{\mu v} z_1^{-\mu} z_2^{-v}}{\sum_{\mu=0}^{n_1} \sum_{v=0}^{n_2} c_{\mu v} z_1^{-\mu} z_2^{-v}} \quad (4.1-4a)$$

oder

$$H(z_1, z_2) = \frac{\underline{z_1}^T \underline{B} \underline{z_2}}{\underline{z_1}^T \underline{C} \underline{z_2}} \quad (4.1-4b)$$

$$\text{mit } \underline{z_\rho}^T = \left[z_\rho^{\alpha_\rho}, z_\rho^{\alpha_\rho-1}, \dots, z_\rho^{-\beta_\rho} \right]$$

mit $\alpha_\rho = r_\rho$ oder 0, $\beta_\rho = m_\rho$ oder n_ρ

für $\rho \in \{1, 2\}$

und der Nennerkoeffizientenmatrix

$$\underline{C} = \begin{bmatrix} c_{00} & c_{01} & \dots & c_{0n_2} \\ c_{10} & c_{11} & \dots & c_{1n_2} \\ \vdots & \vdots & & \vdots \\ c_{n_1 0} & c_{n_1 1} & \dots & c_{n_1 n_2} \end{bmatrix} \quad (4.1-5)$$

Die Zählerkoeffizientenmatrix $\underline{B}$ ist mit (3.1-4) definiert.

- Halbebenenfilter für zeilenweise Verarbeitung in ++ Rekursionsrichtung

$$\begin{aligned} y(k, l) &= \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} u(k-\mu, l-v) \\ &\quad - \sum_{\mu=1}^{n_1} c_{\mu 0} y(k-\mu, l) - \sum_{\mu=-q_1}^{n_1} \sum_{v=1}^{n_2} c_{\mu v} y(k-\mu, l-v) \end{aligned} \quad (4.1-6)$$

$$\begin{aligned} \text{und} \\ H(z_1, z_2) &= \frac{\sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} z_1^{-\mu} z_2^{-v}}{\sum_{\mu=0}^{n_1} c_{\mu 0} z_1^{-\mu} + \sum_{\mu=-q_1}^{n_1} \sum_{v=1}^{n_2} c_{\mu v} z_1^{-\mu} z_2^{-v}} \end{aligned} \quad (4.1-7)$$

mit $c_{00} = 1$.

In diesem Fall hat die Nennerkoeffizientenmatrix $\underline{C}$ in der alternativen Vektorschreibweise nach (4.1-4b) folgendes Aussehen:

$$\underline{C} = [c_{\mu\nu}] = \begin{bmatrix} 0 & c_{-q_1 1} & \dots & c_{-q_1 n_2} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & c_{-1 1} & & c_{-1 n_2} \\ c_{00} & c_{01} & \dots & c_{0 n_2} \\ c_{10} & c_{11} & \dots & c_{1 n_2} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ c_{n_1 0} & c_{n_1 1} & \dots & c_{n_1 n_2} \end{bmatrix} \quad (4.1-8)$$

Für ein symmetrisches Halbebenefilter ist $c_{\mu 0} = 0$ für $\mu = 1, 2, \dots, n_1$.

Im nächsten Abschnitt sollen die Begriffe Verarbeitungsreihenfolge und Rekursionsrichtung so definiert werden, wie wir sie hier verwenden werden. In der Literatur gibt es dazu keine einheitliche Definition.

4.1.2 Verarbeitungsreihenfolge, Rekursionsrichtung

Unter Verarbeitungsreihenfolge soll im folgenden die Reihenfolge verstanden werden, in welcher das Ausgangssignal $y(k, l)$ aus der Differenzengleichung (4.1-3) bzw. (4.1-6) berechnet wird.

In Bild 4.1 sind 4 mögliche Verarbeitungsreihenfolgen angegeben. Je nach Verarbeitungsreihenfolge können dabei die bereits berechneten Ausgangswerte $y(k,l)$ verschiedene Punkte in der (k,l) -Ebene sein.

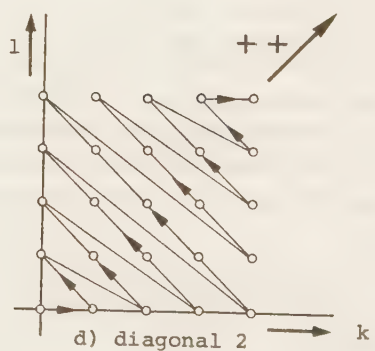
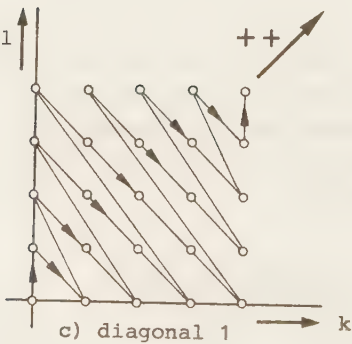
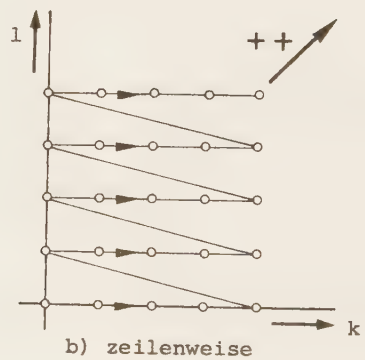
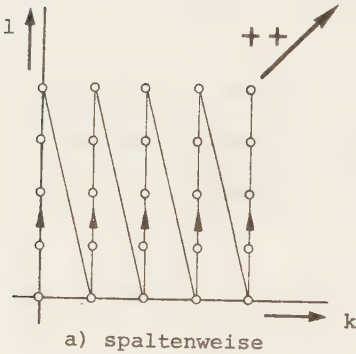


Bild 4.1: Unterschiedliche Verarbeitungsreihenfolge
bei gleicher Rekursionsrichtung

Bei der Verarbeitung ist das Eingangsbild im allgemeinen zeilenweise auf einem Sekundärspeicher abgespeichert. Deshalb ist es zweckmäßig, zur Optimierung der Ein-/Ausgabezeit eine zeilenweise Verarbeitung anzuwenden.

Demgegenüber gibt die Rekursionsrichtung an, in welcher Ecke des 2-D Eingangssignals die Berechnung beginnt und in welcher sie endet.

Definition:

Unter **R e k u r s i o n s r i c h t u n g** wird hier die Richtung des Vektors verstanden, der den ersten Wert des Eingangssignals $u(k,1)$ mit dem letzten zur Berechnung verwendeten Wert des Eingangssignals verbindet.

Je nachdem, ob dabei die Variablen k, l in positiver oder negativer Richtung durchlaufen werden, unterscheidet man 4 verschiedene Rekursionsrichtungen. Dies ist in Bild 4.2 dargestellt.

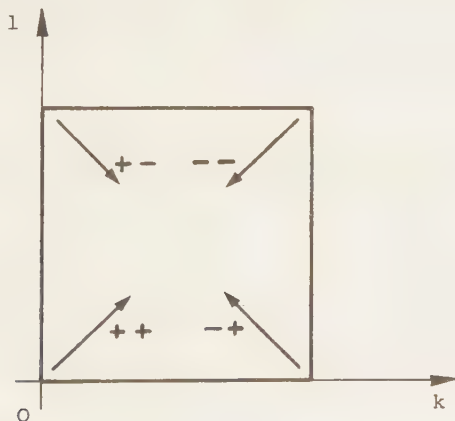


Bild 4.2: Mögliche Rekursionsrichtungen bei einem endlichen Eingangssignal

Der Unterschied zwischen Verarbeitungsreihenfolge und Rekursionsrichtung ist aus Bild 4.1 ersichtlich. Trotz verschiedener Verarbeitungsreihenfolge liegt die gleiche Rekursionsrichtung (++) vor. Die Definition der Kausalität bezieht sich stets auf die Rekursionsrichtung ++.

4.1.3 Kausalität, Semikausalität

Im Gegensatz zur eindimensionalen Signalverarbeitung, bei der in den meisten Fällen zeitabhängige Signale verarbeitet werden, spielt in der zweidimensionalen Signalverarbeitung der Begriff der Kausalität eine untergeordnete Rolle, da es sich hier um Signale handelt, bei denen mindestens eine Variable zeitunabhängig ist. Abgesehen von der bisher kaum möglichen digitalen Echtzeitverarbeitung von zweidimensionalen Signalen liegt das Eingangssignal vollständig vor und muß aus Speicherplatzgründen notwendigerweise in beiden Variablen endlich sein. Da es sich dann bei beiden Variablen um Ortsvariablen handelt, muß Kausalität nicht vorausgesetzt werden. Es kann im allgemeinen auf jeden beliebigen Punkt zugegriffen werden.

Wir behandeln noch einmal den Begriff der Kausalität, für die jetzt betrachteten rekursiven Filter und definieren semikausale Systeme.

Wie in Abschnitt 2.3.1 bereits erwähnt wurde, ist ein zweidimensionales Filter *k a u s a l*, wenn dessen Impulsantwort $h_0(k, l)$ nur im 1. Quadranten der (k, l) -Ebene ($k \geq 0, l \geq 0$) nichtverschwindende Werte annimmt. Ein derartiges Filter wird deshalb Viertelebenenfilter genannt.

Für ein rekursives kausales Filter ist in (4.1-1) und (4.1-2) $q_1 = q_2 = r_1 = r_2 = 0$. Die Kausalität ist dabei unabhängig von der Verarbeitungsreihenfolge.

Von Bedeutung sind außerdem semikausale rekursive Filter.

Definition:

Ein Filter heißt **semikausal**, wenn in (4.1-1) und (4.1-2) gilt:

- a) Verarbeitungsreihenfolge spaltenweise oder diagonal
1 oder 2

$$q_1 = r_1 = 0 \quad \text{und} \quad c_{0v}, b_{0v} = 0 \quad \text{für } v < 0$$

- b) Verarbeitungsreihenfolge zeilenweise

$$q_2 = r_2 = 0 \quad \text{und} \quad c_{\mu 0}, b_{\mu 0} = 0 \quad \text{für } \mu < 0.$$

Die Impulsantwort $h_0(k, l)$ eines derartigen Filters ist dabei entweder auf den 1. und 4. Quadranten (bei a) oder auf den 1. und 2. Quadranten der (k, l) -Ebene (bei b) beschränkt. Für die zeilenweise Verarbeitung in ++ Rekursionsrichtung ist für ein semikausales rekursives Filter in Bild 4.3 der Bereich angegeben, in dem die Impulsantwort $h_0(k, l)$ nichtverschwindende Werte besitzt. Dieser hängt von der Form der Koeffizientenmatrizen $\underline{B}=[b_{\mu v}]$ und $\underline{C}=[c_{\mu v}]$ ab.

Im gezeigten Fall ist das Zählerpolynom kausal, d.h. $\underline{B}=[b_{\mu v}] \quad \mu=0, 1, \dots, m_1, \quad v=0, 1, \dots, m_2$ und das Nennerpolynom semikausal, also $\underline{C}=[c_{\mu v}] \quad \mu=-q_1, -q_1+1, \dots, n_1$
 $v=0, 1, \dots, n_2 \quad c_{\mu 0} = 0 \quad \text{für } \mu < 0$

Der Winkel φ ist bestimmt durch

$$\varphi = \arctan \frac{1}{q_1} \quad (4.1-9)$$

Ein derartiges Filter wird auch als **unsymmetrisches Halbebenenfilter** bezeichnet
 [42, 54].

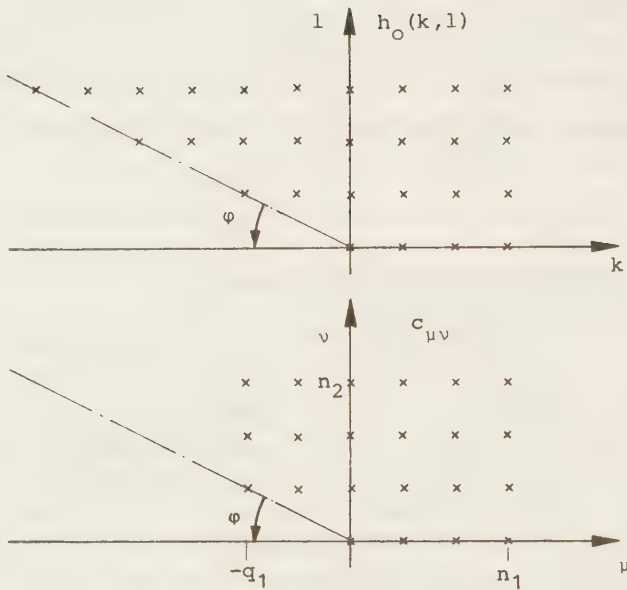


Bild 4.3: Zusammenhang zwischen dem Bereich der Impulsantwort und der Form der Koeffizientenmaske beim unsymmetrischen Halbebenenfilter

Sinnvoller ist es, anstelle des Begriffs der Kausalität den der rekursiven Berechenbarkeit oder auch Rekursierbarkeit einzuführen.

4.1.4 Rekursierbarkeit, Anfangsbedingungen

Definition:

Ein durch eine Differenzengleichung (4.1-1) beschriebenes Filter heißt **r e k u r s i v b e r e c h e n b a r** oder **r e k u r s i e r b a r**, wenn sich jeder Ausgangswert aus den Eingangssignalwerten und bereits berechneten Ausgangssignalwerten bestimmen läßt.

Es gilt:

Ein 2-D rekursives digitales Filter mit der Übertragungsfunktion $H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)}$ ist dann und nur dann rekursiv berechenbar, wenn

$$H_0(z_1, z_2) = \frac{1}{C(z_1, z_2)}$$

bzgl. der genannten Verarbeitungsreihenfolge kausal oder semikausal ist.

Dabei wird vorausgesetzt, daß die Polynome $B(z_1, z_2)$ und $C(z_1, z_2)$ relativ prim zueinander sind, d.h. keine gemeinsamen Faktoren besitzen.

Die einzigen rekursiv berechenbaren Filter (bzgl. ++ Rekursionsrichtung) werden durch die Differenzengleichung (4.1-3) oder (4.1-6) beschrieben.

Zur rekursiven Berechnung der Differenzengleichung ist neben der Forderung der Kausalität bei Viertelebenenfiltern bzw. der Semikausalität bei Halbebenenfiltern noch erforderlich, daß für $y(k, l)$ bestimmte Anfangsbedingungen vorgeschrieben werden.

Für die ++ Rekursion sind bei einem kausalem Viertelebenenfilter die Bereiche, in denen Anfangswerte für das Ausgangssignal $y(k, l)$ vorgeschrieben werden, in Bild 4.4 angegeben.

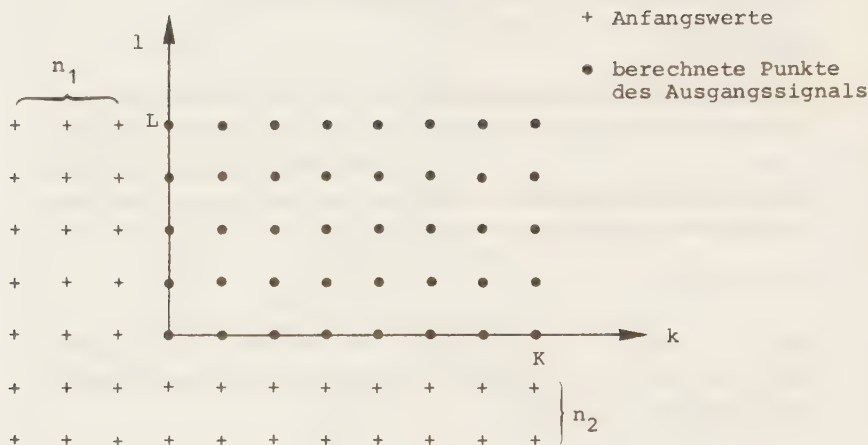


Bild 4.4: Bereiche beim Viertelebenenfilter, in denen Anfangswerte für $y(k, l)$ vorzuschreiben sind

Komplizierter wird dies bei einem semikausalen Halbebenenfilter. Unter der Voraussetzung eines kausalen Eingangssignals und eines kausalen Zählerpolynoms (d.h. in (4.1-6) gelte $r_1=r_2=0$) ist dies in Bild 4.5 dargestellt.

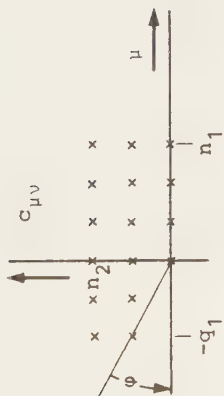
Für lineare Systeme müssen die Anfangswerte stets identisch Null sein [3].

Ein wesentlicher Nachteil für die Filterung mit einem Halbebenenfilter ist aus Bild 4.5 ersichtlich. Wird das Ausgangsbild $y(k, l)$ in dem Ausschnitt $k=0, 1, \dots, K$; $l=0, 1, \dots, L$ gewünscht und ist

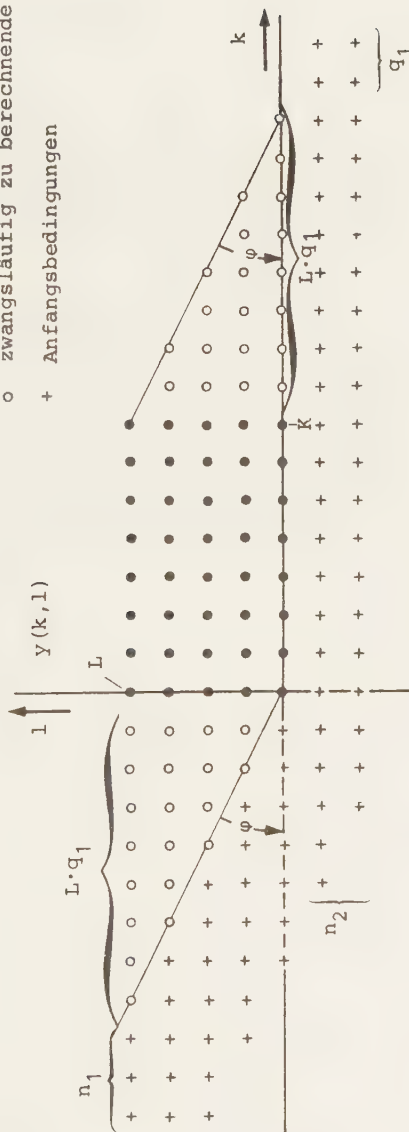
$$\underline{c} = [c_{\mu\nu}] \quad \nu = 0, 1, \dots, n_2$$

$$\mu = -q_1, -q_1+1, \dots, n_1,$$

a)



b)



- gewünschte Ausgangssignalwerte
- zwangsläufig zu berechnende Punkte
- + Anfangsbedingungen

Bild 4.5: a) Form der Nennerkoeffizientenmaske

b) Bereiche beim unsymmetrischen Halbebenenfilter in denen für das Ausgangssignal $y(k, l)$ Anfangswerte vorzuschreiben sind.

so müssen zusätzlich

$$a = (L+1) \cdot L q_1 \quad (4.1-10)$$

Ausgangswerte berechnet werden. D.h. für $K=L=N$ müssen ca. q_1 mal mehr Werte berechnet werden als beim Viertelebenenfilter. Dieser Nachteil kann vermieden werden, wenn am rechten und linken Rand des Ausgangsbildes Randbedingungen vorgeschrieben werden (Bild 4.6).

Dadurch ergibt sich jedoch ein ortsvariantes Filter, das sich durch die Differenzengleichung (4.1-6) nur noch beschreiben läßt, wenn zusätzlich Randbedingungen für $y(k,l)$ angegeben werden. Die Beschreibung durch eine Übertragungsfunktion nach (4.1-7) ist dabei nicht mehr möglich. Dagegen kann die Anordnung als ein eindimensionales System mit N Ein- und Ausgängen aufgefaßt werden [2].

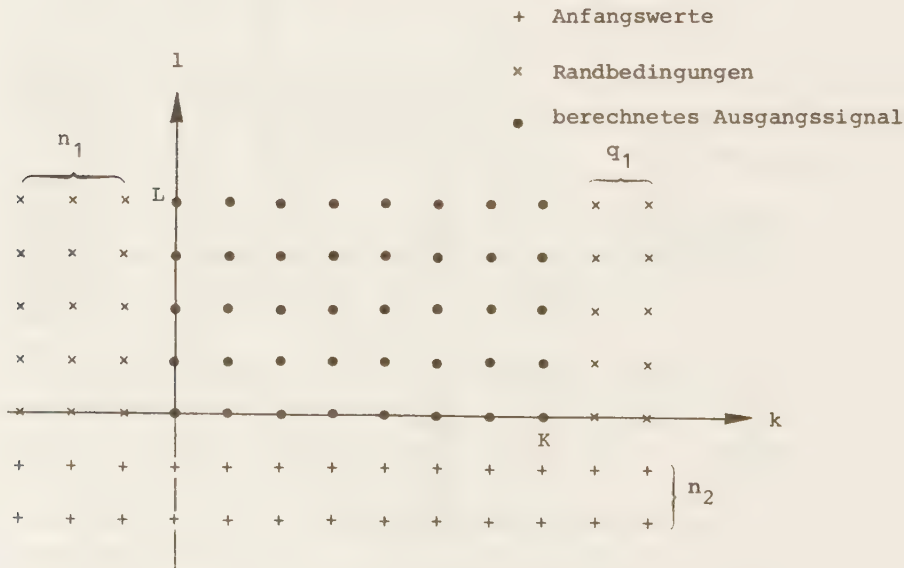


Bild 4.6: Bereiche, in denen Anfangswerte und Randwerte vorgeschrieben werden.

4.2 Stabilität und Stabilitätstests

In diesem Abschnitt wird zunächst eine Übersicht über die wichtigsten Sätze und Tests zur Stabilität zweidimensionaler rekursiver Filter gegeben. Anschließend wird in 4.2.2.3 ein neuer numerischer Stabilitätstest vorgestellt, der sich gegenüber den in der Literatur bekannten Verfahren durch Einfachheit in der Implementierung und kurze Rechenzeit auszeichnet.

Wie in 2.3.1 bereits erwähnt wurde, ist ein rekursives Filter stabil, wenn die Impulsantwort absolut summierbar ist. Anstelle der absoluten Summierbarkeit wird manchmal quadratische Summierbarkeit [31] gefordert. Dies ist zwar eine notwendige, aber keine hinreichende Bedingung für BIBO-Stabilität [30]. Für die Verarbeitung endlicher Signalfolgen genügt es jedoch, die quadratische Summierbarkeit zu fordern [31]. Wird das Ausgangssignal $y(k,l)$ nur in einem endlichen Ausschnitt $k=0,1,\dots,K$; $l=0,1,\dots,L$ berechnet und ist das Eingangssignal $u(k,l)$ ebenfalls auf diesen Ausschnitt begrenzt, so ist lediglich zu fordern, daß die Impulsantwort $h_0(k,l)$ im Bereich $-K \leq k \leq K$, $-L \leq l \leq L$ endlich ist und ein gewünschtes Verhalten besitzt. Um jedoch unabhängig von der Größe des zu berechnenden Ausgangsbildes zu sein, wird im folgenden stets Stabilität vorausgesetzt.

Alle Sätze zur Stabilität beziehen sich ausschließlich auf die Übertragungsfunktion

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)}$$

des rekursiven Systems. Es wird vorausgesetzt, daß Zähler und Nenner keine gemeinsamen Nullstellen besitzen. Dabei genügt es nicht zu fordern, daß die Polynome keine gemeinsamen Faktoren besitzen, also prim zueinander sind, sondern es müssen zudem gemeinsame Nullstellen auf dem Einheitskreis $|z_1| = |z_2| = 1$ ausgeschlossen werden (unwesentliche Singularitäten 2.Art) [30].

Durch die Forderung

$$C(z_1, z_2) \neq 0 \quad \text{für } |z_1| = |z_2| = 1$$

wird dieser Fall ausgeschlossen.

Durch diese Einschränkung handelt es sich bei den nachfolgenden Stabilitätssätzen nur um hinreichende Bedingungen, dennoch sind damit alle praktisch anwendbaren Filter erfaßt, da von der Realisierung eines Filters mit kompensierenden Nullstellen von Zähler- und Nennerpolynom auf dem Einheitskreis schon wegen der endlichen Wortlänge der Koeffizienten und der damit verbundenen Gefahr der Instabilität infolge von Rundungsfehlern stets abzuraten ist.

Unter diesen Voraussetzungen ist allein das Nennerpolynom $C(z_1, z_2)$ für die Stabilität eines rekursiven Systems verantwortlich. Die meisten Stabilitätssätze beziehen sich auf Viertelebenenfilter. Es läßt sich jedoch jedes unsymmetrische Halbebenenfilter endlichen Grades durch lineare Transformation des Indizes μ und ν in (4.1-6) und (4.1-7) in ein Viertelebenenfilter eineindeutig transformieren (siehe Kap. 4.3.2), ohne dadurch die Stabilität des Systems zu beeinflussen [33,34].

4.2.1 Sätze zur Stabilität

Die wichtigsten bekannten Sätze zur Stabilität sind summarisch in dem nachfolgenden Satz zusammengefaßt. Die Beweise finden sich in den zitierten Literaturstellen.

Satz 4.2.1:

Ein rekursives Filter mit der Übertragungsfunktion

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)}$$

ist BIBO stabil, wenn die in Tabelle 4.1 aufgeführten Stabilitätsbedingungen erfüllt sind. Dabei sind alternative Bedingungen durch eine eigene Nummer gekennzeichnet und zusammengefaßt.

Nr.	Bedingung	Literatur
1	$C(z_1, z_2) \neq 0$ für $1 \leq z_1 , 1 \leq z_2 $	28
2a	$C(z_1, z_2) \neq 0$ für $ z_1 = 1, 1 \leq z_2 $	34-37
2b	$C(z_1, \infty) \neq 0$ $1 \leq z_1 $.	
3a	Ist $C(z_1, z_2) \neq 0$ für $ z_1 = z_2 = 1$ reichen nachfolgende Bedingungen für die Stabilität aus	
3b	$C(1, z_2) \neq 0$ für $1 \leq z_2 $	38,39
3c	$C(z_1, \infty) \neq 0$ $1 \leq z_1 $	
4b	$C(\infty, z_2) \neq 0$ für $1 \leq z_2 $	40
4c	$C(z_1, \infty) \neq 0$ $1 \leq z_1 $	
5b	$C(z_1, 1) \neq 0$ für $1 \leq z_1 $	38,39,41
5c	$C(1, z_2) \neq 0$ $1 \leq z_2 $	
6b	$C(z_1, z_2) \neq 0$ für $z_1 = z_2 = z, 1 \leq z $	38
7b	$C(z_1, 1)$ hat keine linearphasigen Anteile für $z_1 = e^{j\Omega_1}$	41
7c	$C(1, z_2)$ hat keine linearphasigen Anteile für $z_2 = e^{j\Omega_2}$.	

Tabelle 4.1: Stabilitätsbedingungen für BIBO-Stabilität

Alle Stabilitätsbedingungen lassen sich aus Bed. 1) ableiten, die besagt, daß das Nennerpolynom $C(z_1, z_2)$ niemals gleichzeitig außerhalb oder auf dem Einheitskreis der z_1 -Ebene und der z_2 -Ebene Nullstellen besitzen darf. Allerdings läßt sich diese Bedingung nur mit erheblichem Rechenaufwand und da auch nicht exakt überprüfen. Deshalb wurden die zitierten alternativen Bedingungen abgeleitet.

Die Bedingungen 2a) und b) stellen den Satz von Huang [34] dar. Ekstrom und Woods [42] haben nachgewiesen, daß dieser Satz auch für unsymmetrische Halbebenenfilter gültig ist.

Da nach (4.1-4a)

$$C(z_1, z_2) = \sum_{\mu=0}^{n_1} \sum_{v=0}^{n_2} c_{\mu v} z_1^{-\mu} z_2^{-v} = \sum_{v=0}^{n_2} C_v(z_1) \cdot z_2^{-v}$$

mit $C_v(z_1) = \sum_{\mu=0}^{n_1} c_{\mu v} \cdot z_1^{-\mu}$ ist,

lassen sich die Bedingungen 2b), 3c), 4c), leicht überprüfen. In diesem Fall muß

$$C_0(z_1) = \sum_{\mu=0}^{n_1} c_{\mu 0} \cdot z_1^{-\mu}$$

ein stabiles Polynom in z_1 sein, d.h. es darf nur Nullstellen im Innern des Einheitskreises der z_1 -Ebene besitzen.

Zur Überprüfung von 2a) beschreibt die Beziehung

$$C(z_1, z_2) = \sum_{v=0}^{n_2} C_v(z_1) \cdot z_2^{-v} = 0, \quad |z_1| = 1$$

eine Abbildung des Randes des Einheitskreises der z_1 -Ebene ($|z_1| = 1$) auf die z_2 -Ebene (Wurzelortskurve [34]).

Bedingung 2a) ist genau dann erfüllt, wenn die Abbildung von $|z_1|=1$ innerhalb des Einheitskreises $|z_2|=1$ in der z_2 -Ebene liegt. Diese Bedingung ist äquivalent zu 3a) und 3b).

Die Bedingung 3a) besagt, daß die Abbildung des Einheitskreises z_1 auf die z_2 -Ebene nicht den Einheitskreis $|z_2|=1$ schneiden darf. Das bedeutet, daß die Fouriertransformierte des Nennerkoeffizientenfeldes keine Nullstellen besitzt. Die Bedingungen (3b-7b) führen auf Stabilitätstests eindimensionaler Polynome, die relativ einfach durchzuführen sind.

Die Bedingungen 3a mit 7b,c sind für alle rekursiven Filter hinreichend, also auch für Halbebenenfilter [38].

Obgleich die Bedingungen 4b,c nur für Viertelebenenfilter gelten, lassen sich diese mit Hilfe der erwähnten Transformation (Kap. 4.3.2) der Indizes μ, ν auf Halbebenenfilter erweitern.

Die meisten Schwierigkeiten bereitet bei allen Stabilitätstests der Nachweis von Bedingung 3a).

4.2.2 Stabilitätstests

In der Literatur beschriebene Stabilitätstests lassen sich grob einteilen in

a) Algebraische Tests

- Koeffizientenbedingungen für Viertelebenenfilter
|28,34,43|
- Marden-Jury-Tafel |44,45|
- Schur-Cohn-Matrix-Test |44,45|

b) Numerische Tests

- α) Cepstrumtest |42,46-48|
- β) Phasentest |31,49|
- γ) Abbildungstests
 - Wurzelortskurven-Test |28,34|
 - Nyquist-Test |38|

Im weiteren werden nur einige der hier aufgeführten Stabilitätstests diskutiert.

Die algebraischen Tests haben den Nachteil, daß sie zwar elegant formuliert werden können, aber für die praktische Anwendung im allgemeinen einen hohen Rechenaufwand erfordern. Für Polynome von Grad $n_1=n_2=n$ steigt dieser Aufwand etwa proportional mit $(n+1)^5$ |31|.

Koeffizientenbedingungen für stabile Viertelebenenfilter mit beliebigen reellen Koeffizienten führen nur für Polynome vom Grad $n_1=n_2=1$ auf einfache Koeffizientenbedingungen.

4.2.2.1 Koeffizientenbedingungen für den 2-D Block 1.Grades

Huang |5,34| und Shanks |28| haben erstmals Bedingungen für die Nennerkoeffizienten abgeleitet, die notwendig und hinreichend für die BIBO-Stabilität eines Filters mit der Übertragungsfunktion

$$H(z_1, z_2) = \frac{1}{\sum_{\mu=0}^1 \sum_{\nu=0}^1 c_{\mu\nu} z_1^{-\mu} z_2^{-\nu}}$$

sind. Mit $c_{00} = 1$ ergibt sich:

$$|c_{10}| < 1 \quad (4.2-1a)$$

$$\frac{c_{01} + c_{11}}{1 + c_{10}} < 1 \quad (4.2-1b)$$

$$\frac{c_{01} - c_{11}}{1 - c_{10}} < 1 \quad (4.2-1c)$$

Dazu äquivalent sind die Bedingungen |50|

$$c_{10} + c_{01} - c_{11} < 1 \quad (4.2-2a)$$

$$c_{10} - c_{01} + c_{11} < 1 \quad (4.2-2b)$$

$$-c_{10} + c_{01} + c_{11} < 1 \quad (4.2-2c)$$

$$-c_{10} - c_{01} - c_{11} < 1 \quad (4.2-2d)$$

Voraussetzung für die Gültigkeit der Ungl. (4.2-1a ...
4.2-1d) sind jedoch reelle Koeffizienten.

Für verschiedene Entwurfsverfahren ist der Fall komplexer Koeffizienten $c_{\mu\nu} \in \mathbb{C}$ interessant. Eine hier vorgenommene Erweiterung des in |43| beschriebenen Verfahrens zur Ableitung der Koeffizientenbedingungen für einen Block 1. Grades mit reellen Koeffizienten unter Verwendung der Marden-Jury-Tafel |56| führt zu den Bedingungen

$$|c_{10}| < 1 \quad (4.2-3a)$$

$$|c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 + 2|c_{01}^* c_{11} - c_{10}| < 1 \quad (4.2-3b)$$

Die Herleitung ist im Anhang 1 angegeben.

Im Fall reeller Koeffizienten geht (4.2-3b) über in

$$c_{11}^2 + c_{01}^2 - c_{10}^2 + 2|c_{01}c_{11} - c_{10}| < 1 \quad .$$

Im Gegensatz zu den Gleichungen (4.2-1a-c) und (4.2-2a-d) hat man hier nur zwei Ungleichungen zu überprüfen.

Für einen rekursiven Block 2.Grades lassen sich ebenfalls notwendige und hinreichende Bedingungen ableiten [51]. Dazu bedient man sich der verallgemeinerten Marden-Jury-Tafel [43] und eines nachfolgenden Sturm-Tests.

Bei einem Systemgrad von $n_1 = n_2 > 2$ sind numerische Tests den algebraischen vorzuziehen.

4.2.2.2 Einige bekannte numerische Testverfahren

Die wohl effektivsten aus der Literatur bekannten numerischen Verfahren zum Stabilitätstest sind der C e p s t r u m - T e s t [40,42,48] und der von Shaw vorgeschlagene P h a - s e n t e s t [31,49].

Der Cepstrum-Test basiert auf dem von Ekstrom und Woods [42] abgeleiteten

Satz 4.2.2:

Ein Filter mit der Übertragungsfunktion

$$H(z_1, z_2) = \frac{1}{C(z_1, z_2)}$$

ist genau dann BIBO-stabil, wenn

$$\hat{c}(\mu, \nu) = Z^{-1}\{\ln C(z_1, z_2)\} \quad (4.2-4)$$

im gleichen Bereich der (μ, ν) -Ebene (Bild 4.7) nichtverschwindende Werte besitzt, wie die zweidimensionale Folge der Nennerkoeffizienten

$$c_{\mu\nu} = Z^{-1}\{C(z_1, z_2)\} . \quad (4.2-5)$$

Dabei ist die komplexe Funktion

$$C(z_1, z_2) = Z\{c_{\mu\nu}\} = |C(z_1, z_2)| \cdot e^{j\phi(z_1, z_2)} \quad (4.2-6)$$

zu logarithmieren, es folgt

$$\hat{C}(z_1, z_2) = \ln C(z_1, z_2) = \ln |C(z_1, z_2)| + j\phi(z_1, z_2) . \quad (4.2-7)$$

Damit das Cepstrum existiert, muß die Phase $\phi(z_1, z_2)$ eine kontinuierliche ungerade periodische Funktion sein [52].

$\hat{C}(\mu, \nu)$ nach (4.2-4) wird komplexes Cepstrum genannt, da $\hat{C}(z_1, z_2)$ in (4.2-7) komplexwertig ist, obgleich für reelle Koeffizienten $c_{\mu\nu}$ das Cepstrum $\hat{C}(\mu, \nu)$ ebenfalls reellwertig ist.

Zur numerischen Durchführung der Cepstrumtests wird die Z-Transformation durch eine zweidimensionale DFT (3.5-1) ersetzt. Da das Cepstrum im allgemeinen unendlich ausgedehnt ist, entstehen Überlappungsfehler (aliasing), die mit grossen Transformationslängen hinreichend klein gehalten werden können. Transformationslängen von $M \times M$ für $M \geq 32$ sind im allgemeinen ausreichend [42].

Zur Berechnung der Phase müssen Phasensprünge um 2π , die auftreten, da die Phase nur im Hauptwertebereich $-\pi \leq \phi \leq \pi$ berechnet wird, eliminiert werden. Dafür ist ein sog. 'phase-

unwrapping'-Algorithmus anzuwenden [55]. Um diesen relativ zeitaufwendigen Algorithmus zu vermeiden, wird von Ekstrom noch ein alternatives Verfahren der Cepstralzerlegung vorgeschlagen [42]. Dazu wird das r e e l l e C e p s t r u m mit $|z_1|=|z_2|=1$

$$\tilde{c}(\mu, \nu) = z^{-1} \{ \ln |C(z_1, z_2)|^2 \} \quad (4.2-9)$$

berechnet.

Viertelebenenfilter

Halbebenenfilter

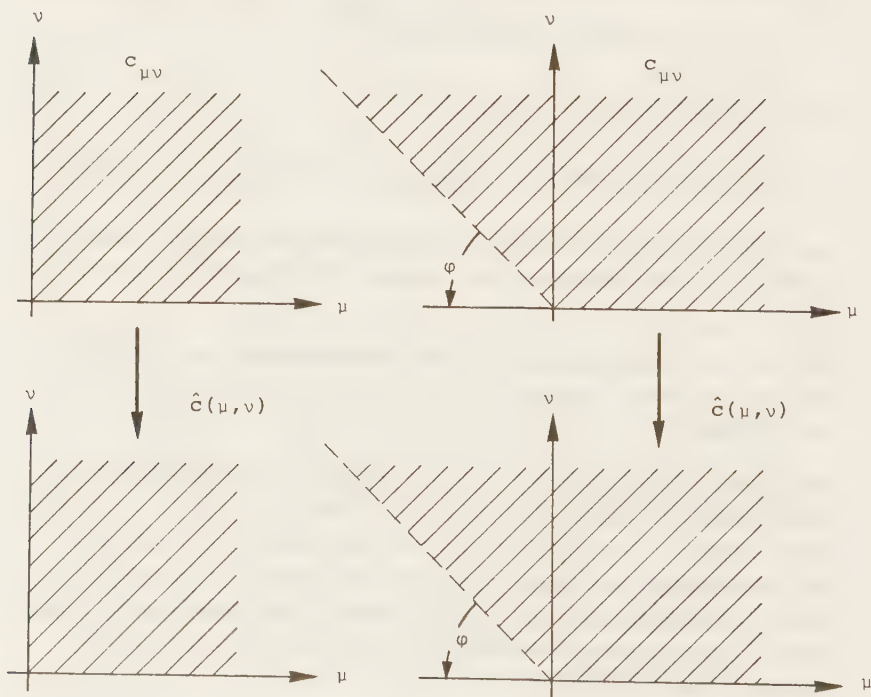


Bild 4.7: Ausdehnung der Nennerkoeffizientenfolge $c_{\mu\nu}$ und des zugehörigen Cepstrum $\hat{c}(\mu, \nu)$ bei BIBO-stabilen Filtern.

$|C(z_1, z_2)|^2$ kann man dabei als das Leistungsspektrum der Autokorrelierten

$$\varphi_{cc}(\mu, \nu) = c(\mu, \nu) ** c(-\mu, -\nu) \quad (4.2-9)$$

interpretieren. Mit Hilfe der Z-Transformation ergibt sich

$$\phi_{cc}(z_1, z_2) = C(z_1, z_2) \cdot C(z_1^{-1}, z_2^{-1}) = |C(z_1, z_2)|^2$$

und durch die Logarithmierung und inverse Z-Transformation

$$\begin{aligned} \tilde{c}(\mu, \nu) &= Z^{-1} \{ \ln C(z_1, z_2) + \ln C(z_1^{-1}, z_2^{-1}) \} \\ &= Z^{-1} \{ \ln C(z_1, z_2) \} + Z^{-1} \{ \ln C(z_1^{-1}, z_2^{-1}) \} \\ &= \hat{c}(\mu, \nu) + \hat{c}(-\mu, -\nu). \end{aligned} \quad (4.2-10)$$

Damit ist die Verknüpfung mit einer Faltung im Ortsbereich (4.2-9) in eine additive Verknüpfung im Cepstralbereich (4.2-10) übergeführt. Nach Satz 4.2.2 läßt sich durch Zerlegung des reellen Cepstrum $\tilde{c}(\mu, \nu)$ der Anteil $\hat{c}(\mu, \nu)$ ermitteln (Spektralfaktorisierung) [42, 53]. Daraus berechnet sich

$$\tilde{c}(\mu, \nu) = Z^{-1} \{ \exp[Z \{ \hat{c}(\mu, \nu) \}] \}. \quad (4.2-11)$$

Ist $c(\mu, \nu) = \tilde{c}(\mu, \nu)$, so ist das Filter stabil, andernfalls instabil.

Das Verfahren der Spektralfaktorisierung läßt sich auch beim Entwurf von Filtern im Frequenzbereich anwenden. Dabei wird das gewünschte Betragsspektrum in einzelne Faktoren zerlegt, die ihrerseits als stabile Filter realisiert werden können [42, 53, 54]. Allerdings führt diese Faktorisierung im allgemeinen nicht auf Faktoren endlichen Grades, weshalb eine Approximation durch Polynome endlichen Grades vorgenommen werden muß [67].

Neben dem relativ großen Rechenaufwand besitzt dieses Verfahren noch einen weiteren signifikanten Nachteil. Analytisch kann das Cepstrum nur dann mit Hilfe der DFT berechnet werden, wenn Bedingung 3a von Satz 4.2.1 erfüllt ist [52]. Dies ist jedoch genau die Bedingung, die überhaupt den hohen Rechenaufwand bedingt, denn die weiteren Bedingungen führen auf eindimensionale Stabilitätstests. Natürlich kann der Cepstraltest ohne Kenntnis davon, ob Bed. 3a erfüllt ist, durchgeführt werden, jedoch ist dieser Test unter diesen Umständen gegebenenfalls nicht mehr schlüssig.

Bereits in [41] wurde festgestellt, daß die 'unwrapped' Phase von $C(z_1, z_2)$ genügt, um die Stabilität eines Filters mit der Übertragungsfunktion $H(z_1, z_2) = \frac{1}{C(z_1, z_2)}$ zu beurteilen. Das Filter ist stabil, wenn die Phase kontinuierlich, ungerade und periodisch ist. Darauf basiert der von Shaw vorgeschlagene Stabilitätstest [31, 49]. Bei diesem Test werden die Bedingungen 3a-c von Satz 4.2.1 überprüft. Mit Hilfe der 2-D FFT und dem 'phaseunwrapping'-Algorithmus von Tribolet [55] wird die Phase an den diskreten Punkten

$$z_1 = e^{j \frac{2\pi k}{K}}, z_2 = e^{j \frac{2\pi l}{L}}, k=0,1,\dots,K-1, l=0,1,\dots,L-1$$

berechnet. Eine n -fache Nullstelle von $C(z_1, z_2)$ auf oder in der nächsten Umgebung des zweidimensionalen Einheitskreises $|z_1|=|z_2|=1$ hat einen Phasensprung der echten Phase um den Wert $n\pi$ zur Folge. Dieses Kriterium dient zur Lokalisierung von möglichen Nullstellen von $C(z_1, z_2)$ auf dem Einheitskreis. Die durch die FFT mit relativ großem Raster ($K=L=32$) lokalisierten möglichen Nullstellen werden innerhalb des gefundenen Intervalls durch ein Nullstellensuchprogramm noch genauer bestimmt und geprüft, ob sie innerhalb oder außerhalb des Einheitskreises liegen.

Im nächsten Abschnitt wird ein neuer Stabilitätstest beschrieben, der den hier erforderlichen relativ rechenintensiven 'phaseunwrapping'-Algorithmus vermeidet.

4.2.2.3 Ein neuer Stabilitätstest

Der hier vorgeschlagene Stabilitätstest basiert ebenfalls auf den Bedingungen 3a-c von Satz 4.2.1, welche für die Stabilität von Viertel- und Halbebenenfilter gültig sind.

Wir betrachten zunächst die Bedingungen 3b,c von Satz 4.2.1:

$$3b) \quad C(1, z_2) \neq 0 \quad \text{für} \quad 1 \leq |z_2|$$

$$3c) \quad C(z_1, \infty) \neq 0 \quad \text{für} \quad 1 \leq |z_1|.$$

Sie führen beide auf Stabilitätstests von Polynomen mit einer Variablen, die z.B. mit der Marden-Jury-Tafel [56] leicht durchgeführt werden können.

Da nach (4.1-7)

$$C(z_1, z_2) = \sum_{\mu=0}^{n_1} c_{\mu 0} z_1^{-\mu} + \sum_{\mu=-q_1}^{n_1} \sum_{v=1}^{n_2} c_{\mu v} z_1^{-\mu} z_2^{-v}$$

gilt, führt Bed. 3b) zu der Forderung

$$C(1, z_2) = C_2(z_2) = \sum_{v=0}^{n_2} c_v z_2^{-v} \neq 0 \quad 1 \leq |z_2| \quad (4.2-12)$$

mit

$$c_v = \begin{cases} \sum_{\mu=0}^{n_1} c_{\mu 0} & \text{für } v = 0 \\ \sum_{\mu=-q_1}^{n_1} c_{\mu v} & \text{für } v = 1, \dots, n_2 \end{cases} \quad (4.2-13)$$

Die Bedingung 3c) ergibt

$$C(z_1, \infty) = C_1(z_1) = \sum_{\mu=0}^{n_1} c_{\mu 0} z_1^{-\mu} \neq 0 \quad 1 \leq |z_1|. \quad (4.2-14)$$

Wesentlicher ist der Test, ob Bedingung

$$3a) \quad C(z_1, z_2) \neq 0 \text{ für } |z_1| = |z_2| = 1$$

erfüllt ist.

Bei einer Nullstelle von $C(z_1, z_2)$ muß sowohl dessen Real- als auch dessen Imaginärteil Null werden. Um das zu kontrollieren, wird zunächst mit Hilfe der zweidimensionalen FFT $C(z_1, z_2)$ an den $M \times M$ diskreten Punkten

$$z_1 = e^{j \frac{2\pi k}{M}}, \quad z_2 = e^{j \frac{2\pi l}{M}} \quad k, l = 0, 1, \dots, M-1$$

berechnet. Damit ergibt sich Real- und Imaginärteil der DFT $C_0(k, l)$. Eine Transformationslänge von $M=32$ oder 64 ist dafür im allgemeinen ausreichend. Bei einer n -fachen Nullstelle von $C(z_1, z_2)$ für $|z_1| = |z_2| = 1$ entsteht ein Phasensprung von $n \cdot \pi$. Bei ungerader Vielfachheit n der Nullstelle tritt stets ein Vorzeichenwechsel von Real- und Imaginärteil von $C(z_1, z_2)$ auf.

Eine mögliche Nullstelle von $C(z_1, z_2)$ kann nur dort auftreten, wo Real- und Imaginärteil von $C_0(k, l)$ entweder das Vorzeichen wechseln oder Null werden. D.h. bei der praktischen Durchführung, daß sie betragsmäßig eine vorgegebene Genauigkeitsschranke ϵ unterschreiten. Letztere ist abhängig von der relativen Genauigkeit des verwendeten Rechners. Bei 32 bit Gleitkommazahlen mit 24 bit Mantisse und 8 bit Exponent ergibt sich eine relative Genauigkeit von ca. 6 Dezimalstellen hinter dem Komma. Zweckmäßigerweise bezieht man dabei ϵ auf den Koeffizienten c_{00} von $C(z_1, z_2)$.

Wird damit eine mögliche Nullstelle von $C(z_1, z_2)$ für

$$z_1 = e^{j\Omega_1}, \quad z_2 = e^{j\Omega_2} \text{ im Intervall } \Omega_1 \in \left\{ \frac{2\pi(k-1)}{M}, \frac{2\pi k}{M} \right\},$$

$\Omega_2 \in \left\{ \frac{2\pi(l-1)}{M}, \frac{2\pi l}{M} \right\}$ lokalisiert, kann für dieses Intervall mit Hilfe eines Optimierungsprogramms eine genaue Bestimmung des Minimums von $C(z_1, z_2)$ durchgeführt werden (z.B.

[57,58]). Für die praktische Anwendung ist dies jedoch überflüssig, da in einem solchen Fall bei Vorgabe einer genügend kleinen Schranke ϵ die Nullstelle von $C(z_1, z_2)$ so dicht am Einheitskreis liegt, daß das Filter bei Koeffizientenrundungsfehlern instabil werden kann. Auf derartige Filter wird man deshalb von vornherein verzichten.

Es ergibt sich das in Bild 4.8 dargestellte Flußdiagramm.

4.2.2.4 Beispiele

Anhand von drei Beispielen

- instabiles Viertelebenenfilter
- stabiles Viertelebenenfilter
- stabiles Halbebenenfilter

soll in diesem Abschnitt die Leistungsfähigkeit des Stabilitätstests von 4.2.2.3 mit dem Cepstraltest nach 4.2.2.2 verglichen werden. Alle gewählten Beispiele sind in der Literatur bekannt.

Beispiel 1:

Wir untersuchen zunächst ein Viertelebenenfilter mit der Übertragungsfunktion $H(z_1, z_2) = \frac{1}{C(z_1, z_2)}$ und der Nennerkoeffizientenmatrix

$$\underline{C} = [c_{\mu\nu}] = \begin{matrix} & \begin{matrix} \nu \rightarrow \\ \begin{bmatrix} 1.0 & -1.200028 & .4000224 \\ -1.00003 & 1.700071 & -.6500509 \\ .4000203 & -.7005488 & .2500439 \end{bmatrix} \end{matrix} \\ \downarrow \mu \\ \begin{bmatrix} 1.0 & -1.200028 & .4000224 \\ -1.00003 & 1.700071 & -.6500509 \\ .4000203 & -.7005488 & .2500439 \end{bmatrix} \end{matrix}$$

Es handelt sich dabei um ein in [28,31] verwendetes Beispiel.

Der in 4.2.2.3 beschriebene Stabilitätstest wird fortan als 'Nullstellensuchtest' bezeichnet. Die Ergebnisausgabe erfolgt in Form eines 'Vorzei-

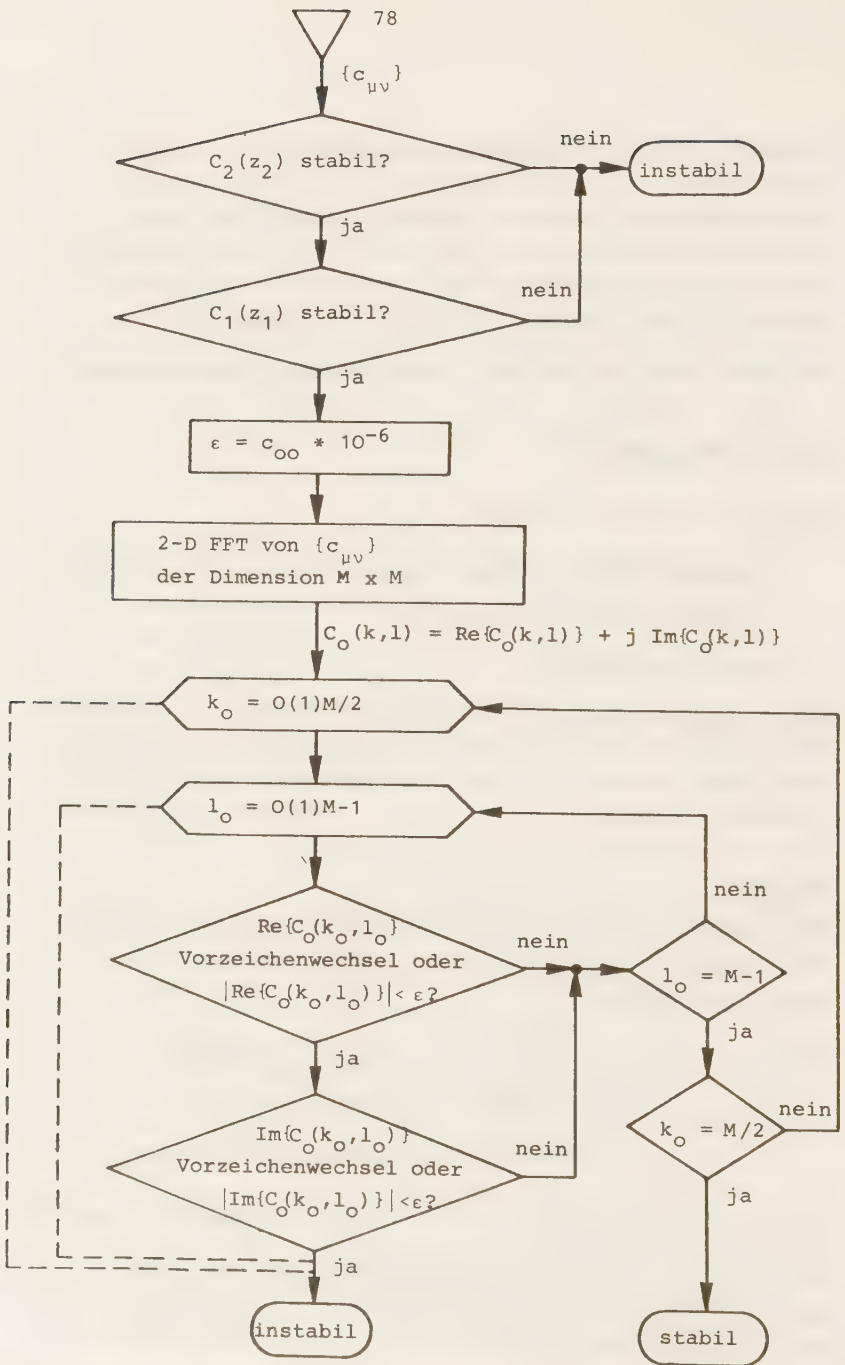


Bild 4.8: Flußdiagramm zum Stabilitätstest von beliebigen rekursiven zweidimensionalen Filtern nach 4.2.2.3

chenverteilungsdiagramms' von Real- und Imaginärteil der 2-D-DFT $C_O(k, l)$ des Nennerpolynoms. Dazu wird abwechselnd das Vorzeichen von Real- und Imaginärteil längs einer Spalte, also in der Reihenfolge

1. Zeile : $\text{Re}\{C_O(0,0)\}, \text{Re}\{C_O(0,1)\}, \dots, \text{Re}\{C_O(0,M-1)\}$
 2. Zeile : $\text{Im}\{C_O(0,0)\}, \text{Im}\{C_O(0,1)\}, \dots, \text{Im}\{C_O(0,M-1)\}$
 3. Zeile : $\text{Re}\{C_O(1,0)\}, \text{Re}\{C_O(1,1)\}, \dots, \text{Re}\{C_O(1,M-1)\}$
 4. Zeile : $\text{Im}\{C_O(1,0)\}, \text{Im}\{C_O(1,1)\}, \dots, \text{Im}\{C_O(1,M-1)\}$
 $\vdots$ $\vdots$ $\vdots$ $\vdots$

N+1. Zeile : $\text{Re}\{C_O(M/2,0)\}, \text{Re}\{C_O(M/2,1)\}, \dots, \text{Re}\{C_O(M/2,M-1)\}$
 N+2. Zeile : $\text{Im}\{C_O(M/2,0)\}, \text{Im}\{C_O(M/2,1)\}, \dots, \text{Im}\{C_O(M/2,M-1)\}$

in folgender Form ausgedruckt (Bild 4.9):

-1 = Vorzeichen negativ
 1 = Vorzeichen positiv
 0 = Betrag kleiner als Genauigkeitsschranke ϵ .

Diese Ausgabe dient nur als Kontrolle und kann weggelassen werden.

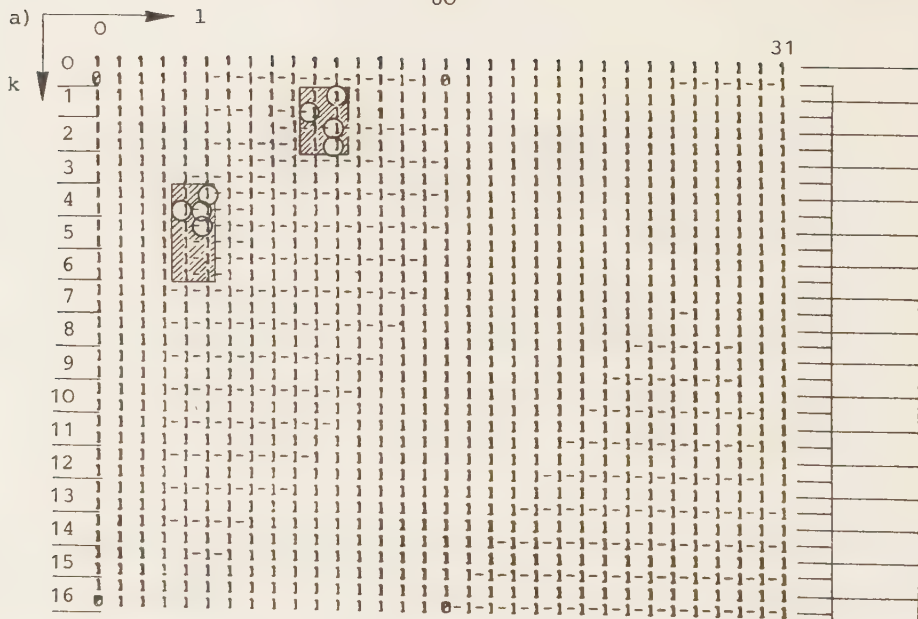
Ausgeführt werden alle Tests mit einer DFT-Länge von 32×32 . Das Filter von Beispiel 1 wird als instabil erkannt. In den Frequenzintervallen

$\Omega_1 \in \{1.767, 1.963\}, \Omega_2 \in \{0.1963, 0.3926\}$ und
 $\Omega_1 \in \{0.785, 0.982\}, \Omega_2 \in \{0.785, 1.178\}$

werden mögliche Nullstellen detektiert.

Beispiel 2: Stabiles Viertelebenenfilter

Beim zweiten Beispiel sei die Koeffizientenmatrix



b)

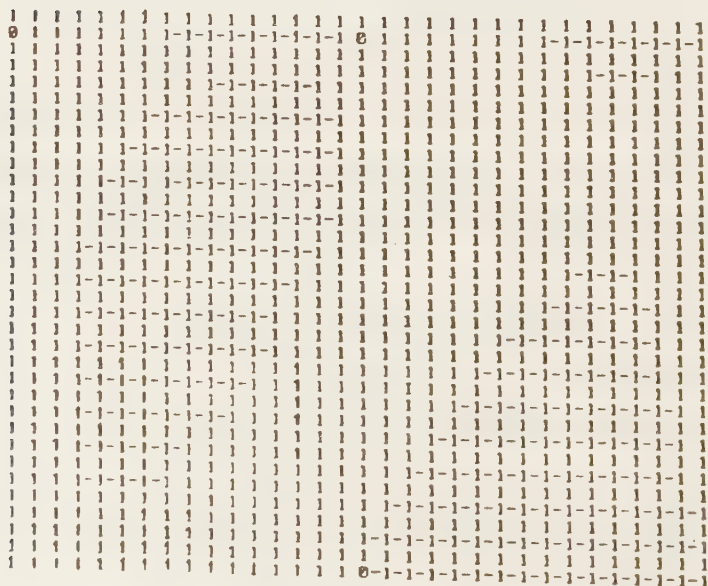
 $Im\{C_O(k, l)\}$ $Re\{C_O(k, l)\}$ 

Bild 4.9: Vorzeichenverteilungsdiagramm von Beispiel 1(a) und Beispiel 2(b)

$$\underline{c} = [c_{\mu\nu}] = \begin{array}{c} \mu \downarrow \begin{array}{c} \nu \rightarrow \\ \begin{bmatrix} 1.00000 & -1.11870 & .349294 \\ -1.00003 & 1.55704 & -.550496 \\ .400020 & -.659803 & .222204 \end{bmatrix} \end{array} \end{array}.$$

Dabei handelt es sich um das nach einer in [41] beschriebenen Methode stabilisierte Viertelebenenfilter von Beispiel 1.

Beispiel 3: unsymmetrisches Halbebenenfilter

Das dritte Beispiel ist ein in [42] angegebenes unsymmetrisches Halbebenenfilter mit der Nennerkoeffizientenmatrix

$$\underline{c} = [c_{\mu\nu}] = \begin{array}{c} \mu \downarrow \begin{array}{c} \nu \rightarrow \\ \begin{bmatrix} 0.0 & 0.000031 & -0.0000172 \\ 0.0 & -0.0343 & -0.0190 \\ 0.0 & 0.0225 & -0.0228 \\ 0.438 & 0.222 & 0.0141 \\ 0.231 & 0.162 & 0.00393 \\ 0.0388 & -0.00356 & -0.0219 \\ -0.0140 & -0.0191 & -0.00789 \end{bmatrix} \end{array} \end{array}.$$

Das Filter wird als stabiles System erkannt.

Sowohl der Cepstraltest als auch der Nullstellensuchtest wurde an einer PDP 11/45 in FORTRAN programmiert. Die Rechenzeit für den Cepstraltest mit reellen Cepstrum (4.2-8) beträgt dabei bei allen drei Beispielen 4,23 sec. Demgegenüber benötigt der Nullstellensuchtest nur 0,86-0,88 sec, also ca. 21% der Zeit des Cepstrumtests.

Im nächsten Abschnitt wird die Realisierung von zweidimensionalen rekursiven Filtern an einem Minicomputer beschrieben. Dabei wird nicht auf spezielle Strukturen (Parallelform, Kaskadenform) eingegangen, sondern es werden zwei prinzipielle Realisierungsmethoden genauer untersucht.

4.3 Realisierung zweidimensionaler rekursiver Filter

In diesem Abschnitt soll die Realisierung von zweidimensionalen rekursiven Filtern unter der Annahme behandelt werden, daß

- das Filter in der direkten Form realisiert wird,
- die Realisierung an einem Minicomputer in Gleitkommaarithmetik erfolgt, wobei idealisierend angenommen wird, daß die Wortlänge hinreichend groß ist, um Fehler infolge endlicher Wortlänge zu vermeiden und
- die Verarbeitung in ++ Rekursionsrichtung (Bild 4.2) durchgeführt wird.

Da $Y(z_1, z_2) = H(z_1, z_2) \cdot U(z_1, z_2) =$

$$= \underbrace{H(z_1, z_2) \cdot z_1^{-r_1} z_2^{-r_2}}_{H_K(z_1, z_2)} \cdot U(z_1, z_2) z_1^{r_1} z_2^{r_2}$$

gilt, kann die Filterung mit einem rekursiven System

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)}$$

mit nichtkausalem Zählerpolynom

$$B(z_1, z_2) = \sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} z_1^{-\mu} z_2^{-v}$$

bei verschobenem Eingangssignal

$u(k, l) \rightarrow u(k+r_1, l+r_2)$ stets durch ein Filter

$$H_K(z_1, z_2) = \frac{B_K(z_1, z_2)}{C(z_1, z_2)} \quad \text{mit kausalem Zählerpolynom}$$

$$B_K(z_1, z_2) = \sum_{\mu=0}^{m_1+r_1} \sum_{v=0}^{m_2+r_2} b_{\mu-r_1, v-r_2} z_1^{-\mu} z_2^{-v}$$

ausgeführt werden.

Der nächste Abschnitt behandelt die Realisierung von Übertragungsfunktionen nichtkausaler Systeme (4.1-2) und von Systemen mit linearer Phase.

4.3.1 Realisierung von nichtkausalen und linearphasigen rekursiven Systemen

Hier werden zwei prinzipiell verschiedene Probleme gemeinsam behandelt:

1. Realisierung eines nichtkausalen Systems mit der Übertragungsfunktion

$$H(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)} = \frac{\sum_{\mu=-r_1}^{m_1} \sum_{v=-r_2}^{m_2} b_{\mu v} z_1^{-\mu} z_2^{-v}}{\sum_{\mu=-q_1}^{n_1} \sum_{v=-q_2}^{n_2} c_{\mu v} z_1^{-\mu} z_2^{-v}} \quad (4.3-1)$$

$$q_1 \neq 0, q_2 \neq 0$$

2. Realisierung eines rekursiven Systems mit linearer Phase bzw. Nullphase

Der Grund dieser Zusammenfassung ist dadurch gegeben, daß beide Punkte auf gleiche Realisierungsmethoden führen:

Wie in Abschnitt 4.1.4 bereits erwähnt wurde, ist ein Filter mit der Übertragungsfunktion nach (4.3-1) nicht mehr rekursiv berechenbar und deshalb auch nicht durch ein Filter, welches nur in einer Rekursionsrichtung angewandt wird, realisierbar. D.h. die Realisierung kann nur durch Kombination von Filtern mit unterschiedlichen Rekursionsrichtungen erfolgen.

Als Kombination ist eine Parallel- oder eine Kaskadenschaltung möglich. Eine Vertauschung der Rekursionsrichtung entspricht einer linearen Transformation (Translation, Rotation) des Eingangssignals und äquivalent dazu der komplexen Frequenzvariablen z_1 und/oder z_2 der Übertragungsfunktion [5, 27, 28].

Ist $H(z_1, z_2)$ die Übertragungsfunktion eines rekursiven Systems für die ++ Rekursionsrichtung (Bild 4.2), so ergibt sich für Rekursion in [27, 53]

$$\begin{aligned} -+ \text{ Richtung} &: H(z_1^{-1}, z_2), \\ +- \text{ Richtung} &: H(z_1, z_2^{-1}), \\ -- \text{ Richtung} &: H(z_1^{-1}, z_2^{-1}). \end{aligned}$$

Abgesehen von separablen Filtern mit der Übertragungsfunktion

$$H(z_1, z_2) = H_0(z_1) \cdot H_0(z_2), \quad (4.3-2)$$

läßt sich mit einem Viertelebenenfilter kein exakt quadrantensymmetrischer Betragsfrequenzgang

$$|\tilde{H}(\Omega_1, \Omega_2)| = |H(e^{j\Omega_1}, e^{j\Omega_2})| = |\tilde{H}(|\Omega_1|, |\Omega_2|)|$$

$$\Omega_v = 2\pi \frac{f_v}{f_{Av}}, \quad f_{Av} = \text{Abtastfrequenz}, \quad v \in \{1, 2\}$$

realisieren [29]. Durch Kombination von zwei Viertelebenenfiltern mit ++ und +- (oder -+) Rekursionsrichtung ist dies jedoch möglich.

Zur Erzeugung von Filtern mit Nullphase müssen zwei entgegengesetzt rekursierende Filter (i.a. ++ und --) kombiniert werden. Um zusätzlich Quadrantensymmetrie des Betragsfrequenzgangs zu erzielen, kombiniert man 4 Viertelebenenfilter mit gleichem Betragsfrequenzgang, aber unterschiedlicher Rekursionsrichtung. Das Gesamtsystem besitzt dann eine quadrantensymmetrische Impulsantwort.

Je nach Kombinationsart unterscheidet man zwischen Nullphasenkaskade mit der Übertragungsfunktion

$$\begin{aligned} H(z_1, z_2) &= H_0(z_1, z_2) \cdot H_0(z_1^{-1}, z_2) \cdot H_0(z_1^{-1}, z_2^{-1}) \cdot H_0(z_1, z_2^{-1}) \\ &= |H_0(z_1, z_2)|^4 \text{ für } z_v = e^{j\Omega_v} \quad v \in \{1, 2\} \end{aligned} \quad (4.3-3a)$$

und der Impulsantwort

$$h(k,l) = h_0(k,l) ** h_0(-k,l) ** h_0(-k,-l) ** h_0(k,-l), \quad (4.3-3b)$$

und P a r a l l e l s c h a l t u n g [5,28] mit der Übertragungsfunktion

$$\begin{aligned} H(z_1, z_2) &= H_0(z_1, z_2) + H_0(z_1^{-1}, z_2) + H_0(z_1^{-1}, z_2^{-1}) + H_0(z_1, z_2^{-1}) \\ &= 2 \operatorname{Re}\{H_0(z_1, z_2) + H_0(z_1^{-1}, z_2)\} \quad \text{für } z_v e^{j\Omega_v}, v \in \{1, 2\} \end{aligned} \quad (4.3-4a)$$

und der Impulsantwort

$$h(k,l) = h_0(k,l) + h_0(-k,l) + h_0(-k,-l) + h_0(k,-l). \quad (4.3-4b)$$

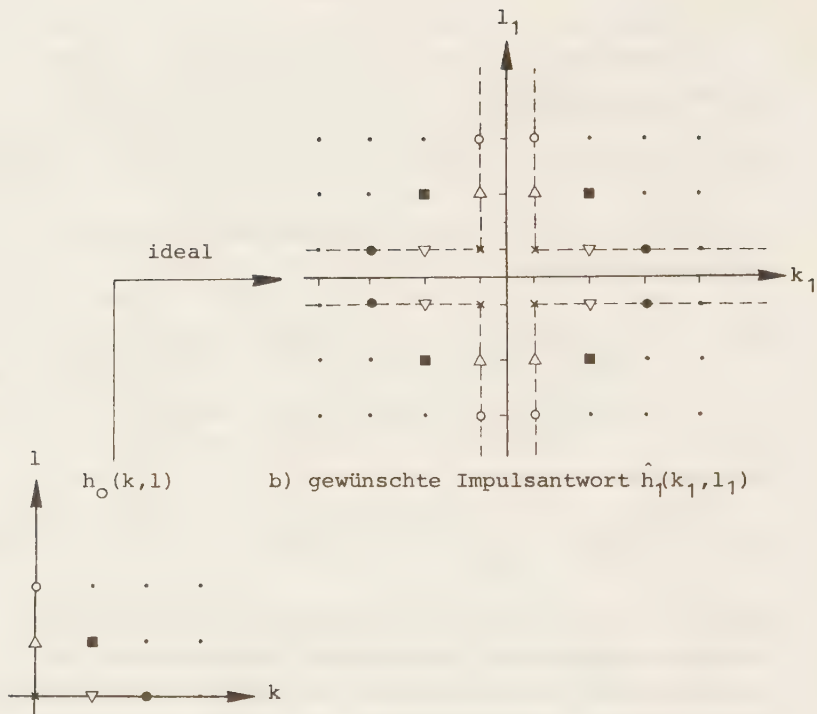
Eine dritte Möglichkeit der Realisierung rekursiver nicht-kausaler Systeme mit linearer Phase wurde erstmals in [87] veröffentlicht. Um die bei der Parallelschaltung auftretende Überlappung der Impulsantworten der Teilfilter (4.3-4b) bei $(k,0), (0,l)$ zu vermeiden, werden die Impulsantworten gegenüber dem Koordinatenursprung in beiden Richtungen um einen halben Bildpunkt verschoben. Es ergibt sich dann die gewünschte Impulsantwort $\hat{h}_1(k_1, l_1)$ nach Bild 4.10b, die jedoch so nicht realisierbar ist.

Realisiert wird stattdessen die um einen halben Bildpunkt in beiden Ortsvariablen verschobene Impulsantwort (Bild 4.10c)

$$\begin{aligned} h_1(k,l) &= \hat{h}_1(k_1 - \frac{1}{2}, l_1 - \frac{1}{2}) = \\ &= h_0(k,l) + h_0(-k-1,l) + h_0(-k-1,-l-1) + h_0(k,-l-1). \end{aligned} \quad (4.3-5b)$$

Die Übertragungsfunktion des realisierten Gesamtsystems

$$\begin{aligned} H_1(z_1, z_2) &= H_0(z_1, z_2) + z_1 H_0(z_1^{-1}, z_2) + z_1 z_2 H_0(z_1^{-1}, z_2^{-1}) + \\ &\quad + z_2 H_0(z_1, z_2^{-1}) \end{aligned} \quad (4.3-5a)$$



a) Impulsantwort des Viertelebenenfilters

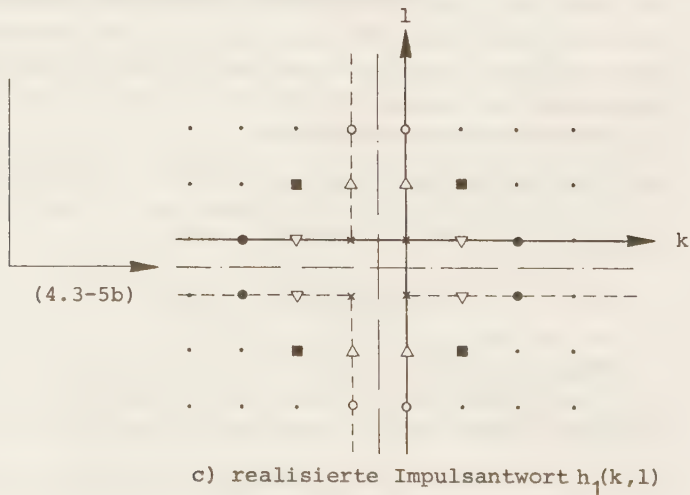


Bild 4.10: Realisierung einer linearphasigen mit quadrantensymmetrischer Impulsantwort durch Parallelschaltung von verschobenen Viertelebenenfiltern

besitzt dann gegenüber dem nichtrealisierbaren nullphasigen System $\hat{H}_1(z_1, z_2)$ mit der Impulsantwort nach Bild 4.10b eine zusätzliche lineare Phase, da

$$H_1(z_1, z_2) = z_1^{-\frac{1}{2}} z_2^{-\frac{1}{2}} \hat{H}_1(z_1, z_2) \text{ ist.}$$

Bei Anwendung dieses Filters wird das Ausgangssignal gegenüber dem Eingangssignal in k und l-Richtung um einen halben Bildpunkt verschoben, was jedoch bei den meisten Anwendungen nicht stört.

Falls das Viertelebenenfilter die Symmetriebedingung $H_0(z_1, z_2) = H_0(z_2, z_1)$ erfüllt, hat das nach (4.3-3a, ..., 4.3-5a) realisierte Gesamtsystem eine oktandensymmetrische Impulsantwort.

Durch geeignete Kombinationen von Halbebenenfiltern lassen sich ebenfalls Systeme mit Nullphase erzeugen [31, 40].

Im nächsten Abschnitt wird gezeigt, wie ein Halbebenenfilter durch ein äquivalentes Viertelebenenfilter realisiert werden kann.

4.3.2 Realisierung eines Halbebenenfilters durch ein Viertelebenenfilter

Wie in 4.2.1 bereits angedeutet wurde, läßt sich jedes Halbebenenfilter umkehrbar eindeutig in ein stabiles Viertelebenenfilter transformieren [33].

Da

$$Y(z_1, z_2) = H(z_1, z_2) \cdot U(z_1, z_2) = \frac{B(z_1, z_2)}{C(z_1, z_2)} \cdot U(z_1, z_2)$$

ist, wobei

$$Y(z_1, z_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} y(k, l) z_1^{-k} z_2^{-l},$$

$$U(z_1, z_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} u(k, l) z_1^{-k} z_2^{-l},$$

$$B(z_1, z_2) = \sum_{\mu=0}^{m_1} \sum_{\nu=0}^{m_2} b_{\mu\nu} z_1^{-\mu} z_2^{-\nu},$$

$$C(z_1, z_2) = \sum_{\mu=0}^{n_1} c_{\mu 0} z_1^{-\mu} + \sum_{\mu=-q_1}^{n_1} \sum_{\nu=1}^{n_2} c_{\mu\nu} z_1^{-\mu} z_2^{-\nu}$$

gilt, ergeben sich nach Anwendung der linearen Transformation

$$T: z_1^{-n} z_2^{-m} \rightarrow z_1^{-(n+q_1 m)} z_2^{-m} \quad (4.3-6)$$

die Transformierten

$$Y_T(z_1, z_2) = T[Y(z_1, z_2)] = \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} y(k, l) z_1^{-(k+l \cdot q_1)} z_2^{-l}$$

$$U_T(z_1, z_2) = \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} u(k, l) z_1^{-(k+l \cdot q_1)} z_2^{-l}$$

$$B_T(z_1, z_2) = \sum_{\nu=0}^{m_2} \sum_{\mu=0}^{m_1} b_{\mu\nu} z_1^{-(\mu+\nu q_1)} z_2^{-\nu}$$

$$C_T(z_1, z_2) = \sum_{\mu=0}^{n_1} c_{\mu 0} z_1^{-\mu} + \sum_{\nu=1}^{n_2} \sum_{\mu=-q_1}^{n_1} c_{\mu\nu} z_1^{-(\mu+\nu q_1)} z_2^{-\nu}$$

und demnach

$$Y_T(z_1, z_2) = \frac{B_T(z_1, z_2)}{C_T(z_1, z_2)} U_T(z_1, z_2) = H_T(z_1, z_2) \cdot U_T(z_1, z_2).$$

Die Transformation nach Gl.(4.3-6) entspricht einer zeilenweisen Verschiebung der Koeffizientenmasken und des Eingangssignals $u(k,l)$ (Bild 4.11). Das Halbebenenfilter $H(z_1, z_2)$ geht dabei in ein Viertelebenenfilter $H_T(z_1, z_2)$ über.

Durch Anwendung der inversen Transformation auf $y_T(k,l)$ erhält man das gewünschte Ausgangssignal $y(k,l)$.

Diese Methode kann auch zum Entwurf von Halbebenenfiltern verwendet werden. Dazu ist das Toleranzschema bzw. die Sollimpulsantwort entsprechend zu transformieren. Es ist ein Viertelebenenfilter zu entwerfen, das die Sollfunktion im Sinne einer vorgegebenen Fehlernorm optimal approximiert [33]. Durch Rücktransformation des so entworfenen Systems erhält man das gewünschte Halbebenenfilter.

In den nächsten beiden Abschnitten werden zwei Realisierungsarten für 2-D rekursive Filter besprochen, die sich sowohl in der Beschreibung als auch bzgl. des Realisierungsaufwands (Kap.4.3.5) unterscheiden.

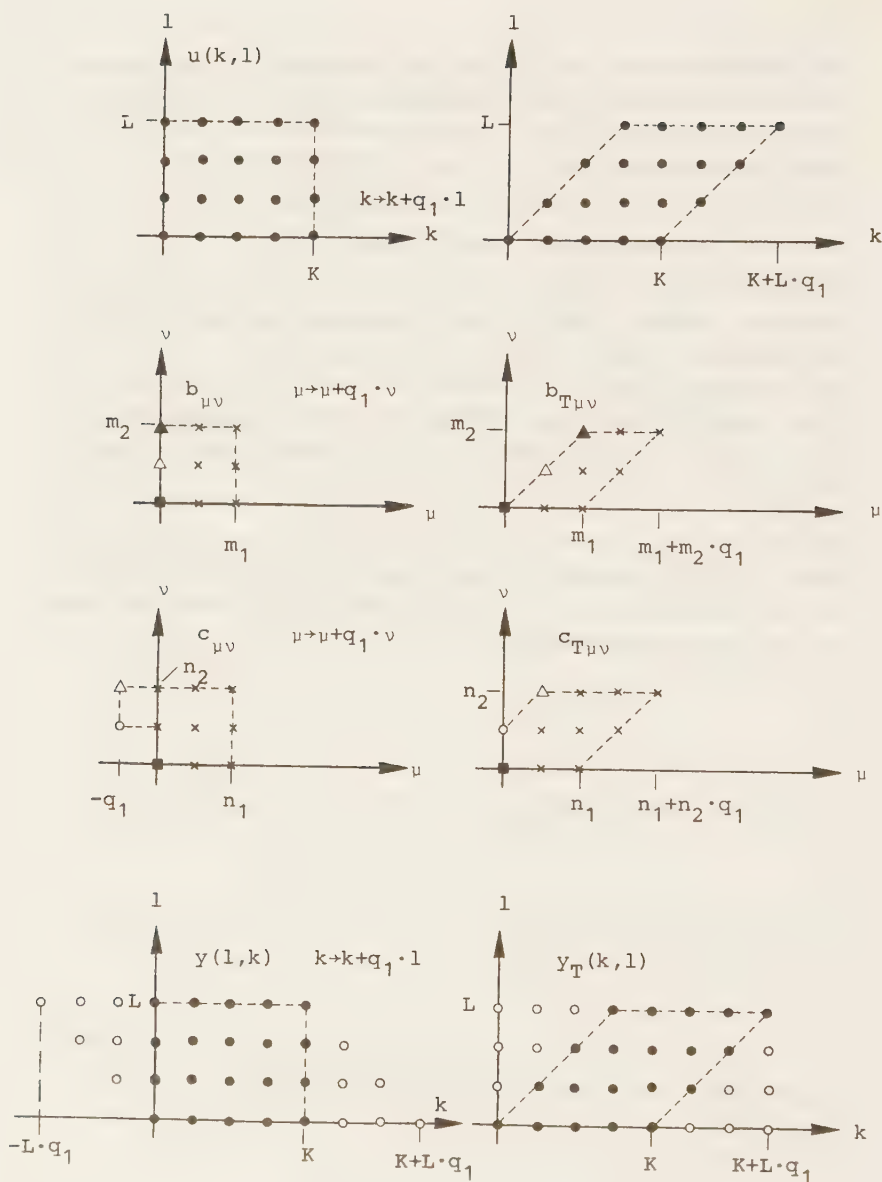


Bild 4.11: Realisierung eines Halbebenenfilters durch ein Viertelebenenfilter am Beispiel

$$m_1 = m_2 = n_1 = n_2 = 2, q_1 = 1$$

4.3.3 Zweidimensionale Realisierung

Für ein Viertelebenenfilter mit der Differenzengleichung (4.1-3) und der Übertragungsfunktion nach (4.1-4a,b) ergibt sich das gewünschte Ausgangssignal durch Verschiebung der Zählerkoeffizientenmaske $\underline{B}$ (3.1-4) über dem Eingangsbild $u(k,l)$ und der Nennerkoeffizientenmaske $\underline{C}$ (4.1-5) über dem Ausgangsbild $y(k,l)$. Für den Fall $r_1=r_2=0$, $m_1=m_2=n_1=2$, $n_2=3$ ist dies in Bild 4.12 für die Berechnung des Ausgangssignalwertes $y(4,6)$ dargestellt.

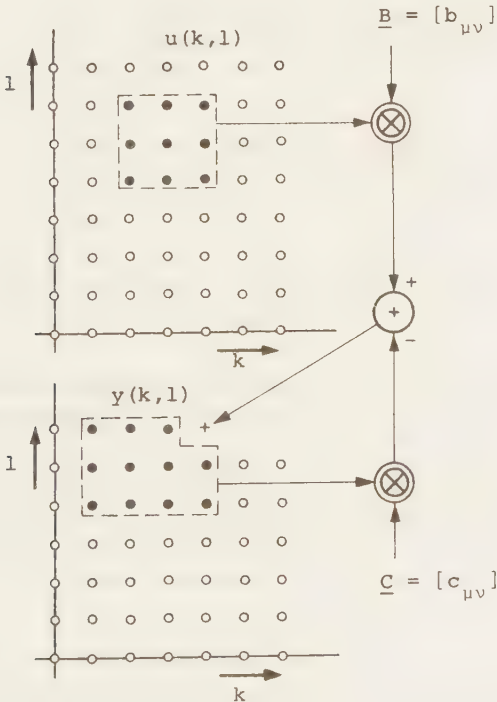


Bild 4.12: Rekursive punktweise Berechnung des Ausgangsbildes

4.3.4 Eindimensionale Realisierung

Ausgehend von der Differenzengleichung (4.1-6) erhält man durch Ausführung der Z-Transformation bzgl. z_1

$$Y_1(z_1) = \frac{1}{C_0(z_1)} \sum_{v=-r_2}^{m_2} B_v(z_1) U_{1-v}(z_1) - \sum_{v=1}^{n_2} C_v(z_1) Y_{1-v}(z_1) \quad (4.3-7)$$

$$\text{mit } B_v(z_1) = \sum_{\mu=-r_1}^{m_1} b_{\mu v} z_1^{-\mu}, \quad v = -r_2, -r_2+1, \dots, m_2 \quad (4.3-8)$$

und

$$C_v(z_1) = \begin{cases} 1 + \sum_{\mu=1}^{n_1} c_{\mu 0} z_1^{-\mu} & v = 0 \\ \sum_{\mu=-q_1}^{n_1} c_{\mu v} z_1^{-\mu} & v = 1, 2, \dots, n_2. \end{cases} \quad (4.3-9)$$

Dabei sind $U_1(z_1)$, $Y_1(z_1)$ die eindimensionalen Z-Transformierten der 1-ten Zeile des Ein- bzw. Ausgangsbildes.

Gl.(4.3-7) gibt eine Vorschrift über die Berechnung einer vollständigen (der 1-ten) Zeile des Ausgangsbildes aus n_2 vorher berechneten Zeilen des Ausgangsbildes und r_2+m_2+1 Zeilen des Eingangsbildes.

$$H_0(z_1) = \frac{1}{C_0(z_1)} \quad (4.3-10)$$

stellt dabei ein eindimensionales rekursives Filter dar. Damit das zweidimensionale Gesamtsystem rekursiv berechenbar ist, muß $H_0(z_1)$ kausal sein. Für ein allgemeines rekursives Filter nach (4.1-6) in direkter Form ergibt sich der

in Bild 4.13 dargestellte Signalflußgraph.

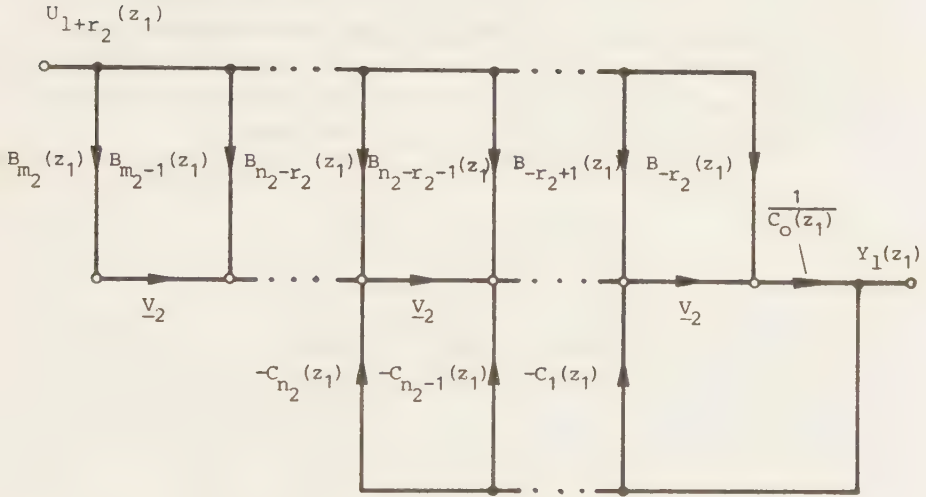


Bild 4.13: Signalflußgraph für ein allg. 2-D rekursives Filter in der direkten Form (1. kanonische Form)

Dabei sind die $B_v(z_1)$, $v = -r_2, \dots, m_2$ 1-D nichtrekursive Filter der Länge m_1+r_1+1 und die $C_v(z_1)$, $v = 1, \dots, n_2$ nichtrekursive Filter der Länge q_1+n_1+1 . Damit das System realisierbar wird, muß man sich nach der rekursiven Filterung mit $H_o(z_1)$ auf einen endlichen Signalausschnitt beschränken. Die Berechnung erfolgt dabei zeilenweise. Das Verzögerungselement V_2 speichert jeweils eine ganze Zeile.

Solange die nichtrekursiven Filter $B_v(z_1)$, $C_v(z_1)$ einschließlich $C_o(z_1)$ kausal sind, läßt sich das Ausgangssignal $y_1(k)=y(k,l)$ und das Eingangssignal $u_1(k)=u(k,l)$ (mit l als Parameter) bzgl. der Ortsvariablen k auf die gleiche "Zeilenlänge" $k=0,1,\dots,N-1$ begrenzen. Die Zustandsvariablen des rekursiven Systems $H_o(z_1)$ (4.3-10) müssen dabei nach jeder Zeile gelöscht werden.

Für den Sonderfall, daß $C_0(z_1) = \text{const}$ ist, bleibt die Impulsantwort und damit das Ausgangssignal bzgl. k begrenzt. D.h. das System verhält sich bzgl. dieser Ortsvariablen als nichtrekursives System. Für $r_1=r_2=q_2=0$ und $q_1=n_1=m_1=n$ in (4.1-6) ist in Bild 4.14 der Bereich angegeben, in dem die Impulsantwort nichtverschwindende Werte besitzt.

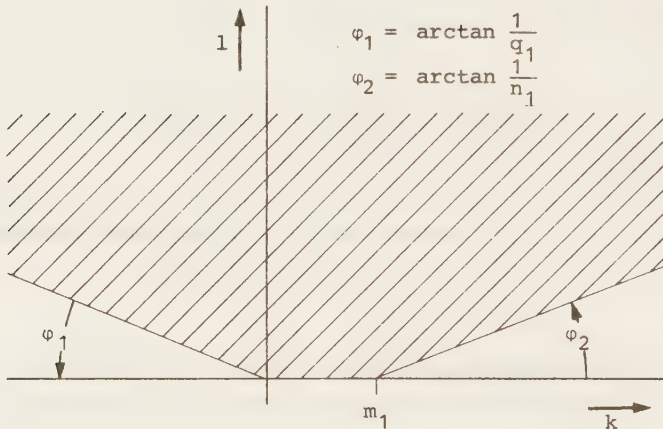


Bild 4.14: Bereich für die Begrenzung der Impulsantwort bei einem symmetrischen Halbebenenfilter mit $r_1=r_2=q_2=0$ und $q_1=n_1=m_1=n$ in (4.1-6)

Die 1-D-Filterung mit den im allgemeinen nichtkausalen nichtrekursiven Filtern kann stets durch eine Filterung mit einem entsprechend verschobenen kausalen Filter und anschließender Rückverschiebung ersetzt werden. Das eindimensionale rekursive Filter $H_0(z_1)$ muß nicht notwendigerweise kausal sein, sondern kann auch als eindimensionales Nullphasenfilter realisiert werden [59].

Für ein kausales ortsinvariantes System wird nun eine alternative Beschreibung zu (4.3-7) angegeben, wobei das als kausal angenommene Eingangssignal $N \times N$ Bildpunkte umfaßt und das Ausgangssignal auf den gleichen Ausschnitt begrenzt wird.

$$\text{Mit } \underline{u}(1) = \underline{u}_1 = \left[u_1(0), \dots, u_1(N-1) \right]^T \quad (4.3-11)$$

$$\underline{y}(1) = \underline{y}_1 = \left[y_1(0), \dots, y_1(N-1) \right]^T$$

ergibt sich für $m_2=n_2$ und $r_2=0$ die zu (4.1-6) äquivalente Differenzengleichung in Matrixschreibweise

$$\underline{y}(1) = \underline{C}_0^{-1} \left[\sum_{v=0}^{n_2} \underline{B}_v \cdot \underline{u}(1-v) - \sum_{v=1}^{n_2} \underline{C}_v \cdot \underline{y}(1-v) \right] \quad (4.3-12)$$

Im Fall kausaler nichtrekursiver Systeme $B_v(z_1)$, $C_v(z_1)$ $v=0,1,\dots, n_2$ werden die $N \times N$ Matrizen $\underline{B}_v, \underline{C}_v$ untere Töplitzmatrizen, wobei für $m_1=n_1$

$$\underline{B}_v = \begin{bmatrix} b_{0v} & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ b_{n_1 v} & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & b_{n_1 v} & \dots & b_{0v} \end{bmatrix} \quad (4.3-13)$$

$$\underline{C}_v = \begin{bmatrix} c_{0v} & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ c_{n_1 v} & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c_{n_1 v} & \dots & c_{0v} \end{bmatrix}$$

$v=0,1,\dots, n_2$ gilt.

Die inverse Matrix $\underline{C}_0^{-1}$ ist dabei ebenfalls wieder eine untere Töplitzmatrix. Die Matrixelemente der ersten Spalte sind identisch mit den ersten N Werten der Impulsantwort des rekursiven Systems $H_0(z_1) = \frac{1}{C_0(z_1)}$.

Diese Darstellung entspricht der in [2] angegebenen Beschreibung als System mit N Ein- und Ausgängen (Bild 4.15).

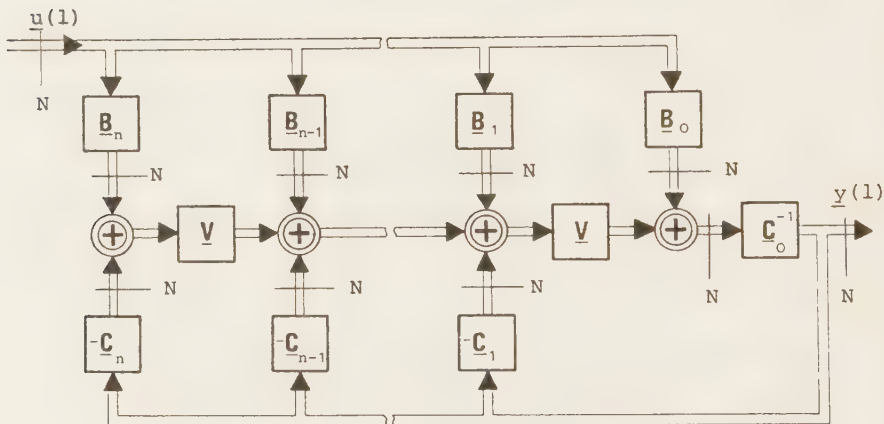


Bild 4.15: Struktur des durch die Differenzengleichung (4.3-12) beschriebenen Systems

Betrachtet sei nun der Sonderfall, daß in (4.3-7) $C_0(z_1)=1$ ist. Aus (4.3-12) folgt dann

$$\underline{y}(1) = \sum_{v=0}^{n_2} \underline{B}_v \underline{u}(1-v) - \sum_{v=1}^{n_2} \underline{C}_v \underline{y}(1-v). \quad (4.3-14)$$

Das System nach (4.3-14) werde nun mit dem Impuls nach (2.3-2)

$$\underline{u}(1) = \underline{y}_0(1) = \begin{cases} 1 & l=0 \\ 0 & \text{sonst} \end{cases} \quad (4.3-15)$$

$$\text{mit } \underline{1} = \overbrace{[1 \ 0 \ \dots \ 0]}^N]^T$$

$$\text{und } \underline{0} = \underbrace{[0 \ 0 \ \dots \ 0]}_N]^T$$

erregt.

Aus (4.3-14) ergibt sich damit als Systemreaktion die "Vektorimpulsantwort"

$$\underline{h}_0(l) = \sum_{v=0}^{n_2} \underline{B}_v \underline{y}_0(l-v) - \sum_{v=1}^{n_2} \underline{C}_v \underline{h}_0(l-v), \quad (4.3-16)$$

für die wegen der Kausalität bzgl. der zweiten Ortsvariablen l $\underline{h}_0(l) = \underline{0}$ für $l < 0$ gilt.

Mit (4.3-16) läßt sich das 2-D Gesamtsystem für den Bildausschnitt von $N \times N$ Werten als System mit $N \times N$ Ein- und Ausgängen beschreiben zu

$$\begin{bmatrix} \underline{y}(0) \\ \underline{y}(1) \\ \vdots \\ \underline{y}(N-1) \end{bmatrix} = \begin{bmatrix} \underline{H}_0 & \underline{0} & \dots & \underline{0} \\ \underline{H}_1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \underline{H}_{N-1} & \cdot & \cdot & \underline{H}_1 \end{bmatrix} \cdot \begin{bmatrix} \underline{u}(0) \\ \underline{u}(1) \\ \vdots \\ \underline{u}(N-1) \end{bmatrix}, \quad (4.3-17)$$

wobei $\underline{0} = [\underline{0} \ , \ \underline{0} \ \dots \ \underline{0}]^T$ ist und für ein bzgl. der ersten Ortsvariablen kausales System

$$\underline{h}_0(l) = \underbrace{\begin{bmatrix} h_0(0,l) & 0 & \dots & 0 \\ h_0(1,l) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 \\ h_0(N-1,l) & \cdot & \cdot & h_0(1,l) h_0(0,l) \end{bmatrix}}_{\underline{H}_1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.3-18)$$

die ersten N Werte der Impulsantwort der l -ten Zeile enthält.

Gl. (4.3-17) beschreibt eine kausale ortsinvariante nichtrekursive Filterung vom Blockgrad $O \mid 4 \mid$.

Vorstehende Überlegungen gelten auch, wenn $C_O(z_1)$ nicht mehr konstant, sondern ein kausales eindimensionales Polynom in z_1 ist. In diesem Fall sind in (4.3-14) die Koeffizientenmatrizen $\underline{B}_v$, $\underline{C}_v$ (4.3-13) durch die Matrizen $\underline{C}_O^{-1} \cdot \underline{B}_v$, $\underline{C}_O^{-1} \cdot \underline{C}_v$ zu ersetzen. Für bzgl. der Ortsvariablen k nichtkausale Teilfilter $B_v(z_1)$, $C_v(z_1)$ $v=0, \dots, n_2$ gilt die Beschreibung nach (4.3-12) ebenfalls, in diesem Fall ist jedoch die Impulsantwort $\underline{h}_O(l)$ im Bereich $-N+1 \leq k \leq N-1$ zu berechnen. Die Matrizen $\underline{H}_l$ für $l=0, \dots, N-1$ in (4.3-17) ergeben sich zu

$$\underline{H}_l = \begin{bmatrix} h_O(0,l) & h_O(-1,l) & \dots & h_O(-N+1,l) \\ h_O(1,l) & \cdot & \cdot & \cdot \\ \vdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ h_O(N-1,l) & \dots & h_O(1,l) & h_O(0,l) \end{bmatrix} \quad (4.3-19)$$

Der Sonderfall $h_O(k,l) = h_O(-k,l)$ führt zu einem nullphasigen Filter in k -Richtung. Dieser Fall läßt sich erreichen, wenn $C_O(z_1)=1$ ist und in (4.3-7 bis (4.3-9) $r_1=m_1=m$, $q_1=n_1=n$ wird.

Außerdem müssen die eindimensionalen nichtrekursiven Systeme $B_v(z_1)$, $C_v(z_1)$ nullphasig sein, das ist der Fall, wenn $b_{\mu v} = b_{-\mu v}$, $\mu=0,1,\dots,m$ und $c_{\mu v} = c_{-\mu v}$ für $\mu=0,1,\dots,n$ gilt.

Randeffekte bei der nichtrekursiven Teilfilterung lassen sich durch ortsvariante eindimensionale Randfilterung (Kap. 3.2.2) weitgehend unterdrücken.

4.3.5 Realisierungsaufwand für die Durchführung der 2-D rekursiven Filterung

Analog zu Abschnitt 3.6 wird hier der Realisierungsaufwand für die Ausführung der rekursiven zweidimensionalen Filterung an einem Minicomputer angegeben.

Dazu werden folgende Annahmen getroffen:

- Begrenzung des Ein- und Ausgangssignals auf $N \times N$ -Werte,
- kausales zweidimensionales System,
- Realisierung in direkter Form (1. kanonische Form),
- Zählergrad = Nennergrad = $n_r \cdot n_r$
- N_B reelle Datenworte im Hauptspeicher,
- zeilenweise Verarbeitung,
- das Eingangssignal ist (zeilenweise) sequentiell auf einem Sekundärspeicher abgelegt,
- das Ausgangssignal wird (zeilenweise) sequentiell auf dem Sekundärspeicher abgespeichert,
- Koeffizientensymmetrien werden nicht berücksichtigt.

Unter diesen Annahmen erhält man sowohl für die Realisierung nach 4.3.3 (Bild 4.12) als auch für die eindimensionale Realisierung nach 4.3.4 (Bild 4.13) den gleichen Aufwand für die arithmetische Berechnung.

Es sind pro Ausgangswert

$2(n_r+1)^2-1$ reelle Multiplikationen,

und $2(n_r+1)^2-2$ reelle Additionen

erforderlich. Mit den Bestimmungsgrößen nach 3.5.1 ergibt sich als Zeit t_A für die arithmetische Berechnung

$$t_A = N^2 \{ [2(n_r+1)^2-1] \cdot T_{\text{mult}} + [2(n_r+1)^2-2] \cdot T_{\text{add}} \}. \quad (4.3-20)$$

Bezüglich des notwendigen Speicheraufwands unterscheiden sich die direkte 2-D-Realisierung nach 4.3.3 und die 1-D-Realisierung nach 4.3.4 jedoch wesentlich. Es ergibt sich als Mindestspeicheraufwand (für die Berechnung einer Zeile) bei

zweidimensionaler Realisierung : $2(n_r+1)$ Zeilen

eindimensionaler Realisierung : n_r+2 Zeilen.

Demnach sind bei der eindimensionalen Realisierung nach Bild 4.13 n_r Zeilen weniger zu speichern als bei der direkten zweidimensionalen Realisierung. (Die notwendigen n_r Zustandsvariablenspeicher für die eindimensionale nichtrekursive Filterung wurden dabei vernachlässigt.) Dieses Verfahren ist deshalb hier vorzuziehen. Bei fest vorgegebenem Speicherplatz N_B können in diesem Fall mehr Zeilen des Ausgangssignals ohne Zugriff auf den Sekundärspeicher berechnet werden als bei der direkten zweidimensionalen Realisierung. Die Zahl der Speicherzugriffe ist bei der

zweidimensionalen Realisierung: $2 \cdot \left\langle \frac{N^2}{N_B - (2n_r+1)N} \right\rangle$,

und bei der

eindimensionalen Realisierung : $2 \cdot \left\langle \frac{N^2}{N_B - (n_r+1)N} \right\rangle$.*)

Damit ergibt sich als Zeit für den Datentransfer vom und zum Sekundärspeicher mit den Bestimmungsgrößen T_{acc} und T_{tran} nach 3.5.1 für die

- zweidimensionale Realisierung

$$t_{IO} = 2 \left\langle \frac{N^2}{N_B - (2n_r+1)N} \right\rangle T_{acc} + 2 N^2 T_{tran} , \quad (4.3-21a)$$

*) $\langle x \rangle$: kleinste ganze Zahl, größer oder gleich x

für die

- eindimensionale Realisierung

$$t_{IO} = 2 \left\langle \frac{N^2}{N_B - (n_r + 1)N} \right\rangle T_{acc} + 2 N^2 T_{tran} . \quad (4.3-21b)$$

Für $N_B = 2^{13}$ mit Bildgröße N als Parameter sind in Bild 4.16 die Zeiten t_A und t_{IO} für die eindimensionale Realisierung in Abhängigkeit von der Filterlänge $n_r + 1$ dargestellt.

4.4 Vergleich des Realisierungsaufwands von zweidimensionalen nichtrekursiven und rekursiven Filterung

Es ist bekannt, daß die Approximation eines vorgegebenen Wunschfrequenzgangs mit 1-D rekursiven Filtern bei gleicher Approximationsgüte auf einen wesentlich niedrigeren Filtergrad führt als vergleichbare nichtrekursive Filter erfordern [1]. Deshalb ergibt sich für die Realisierung mit rekursiven Filtern ein kleinerer Rechen- und Speicherplatzaufwand als bei nichtrekursiven Filtern. Dies gilt auch für den zweidimensionalen Fall. Bei zweidimensionalen nichtrekursiven Filtern sind im allgemeinen ca. 10mal mehr Koeffizienten erforderlich als bei rekursiven zweidimensionalen Filtern [60].

In Bild 4.17 ist für einen Hauptspeicherplatz von $N_B = 2^{13}$ reellen Worten die Gesamtrechnenzeit $t_G = t_A + t_{IO}$ für die nichtrekursive Nullphasenfilterung mit direkter Faltung (3.6-2, 3.6-3) und der FFT-Implementierung (3.6-4, 3.6-5) sowie der rekursiven Filterung (4.3-20, 4.3-21b) in Abhängigkeit von der Filterlänge n dargestellt. Der Parameter N gibt dabei wieder die Dimension $N \times N$ des Ein- und Ausgangsbildes an.

Wir führen einen Vergleich für den besonders interessanten Fall durch, daß eine Verarbeitung durch Filter mit Nullphase und Quadrantensymmetrie bzgl. Betragsfrequenzgang

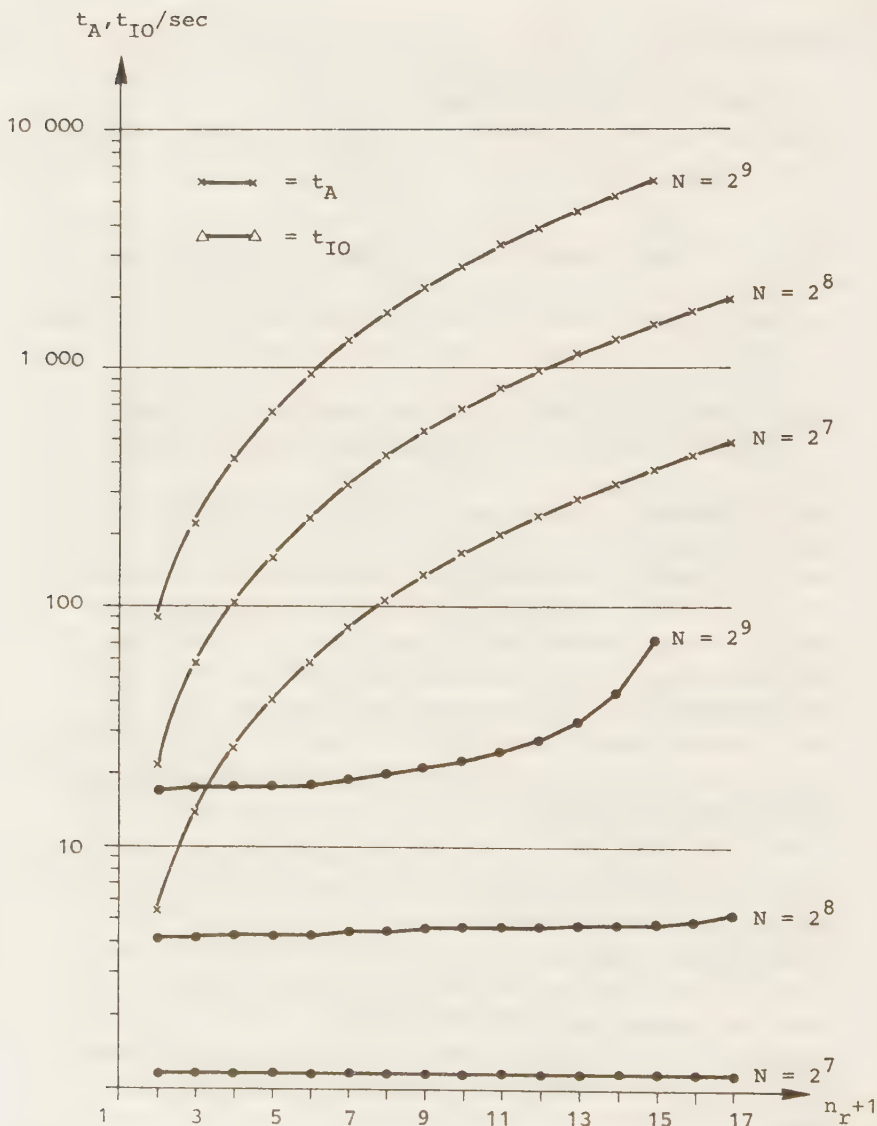


Bild 4.16: Zeit t_A für die arithmetische Berechnung und den Datentransfer t_{IO} bei der rekursiven Filterung in Abhängigkeit von der Filterlänge $n_r + 1$ und der Bildgröße N für $N_B = 2^{13}$

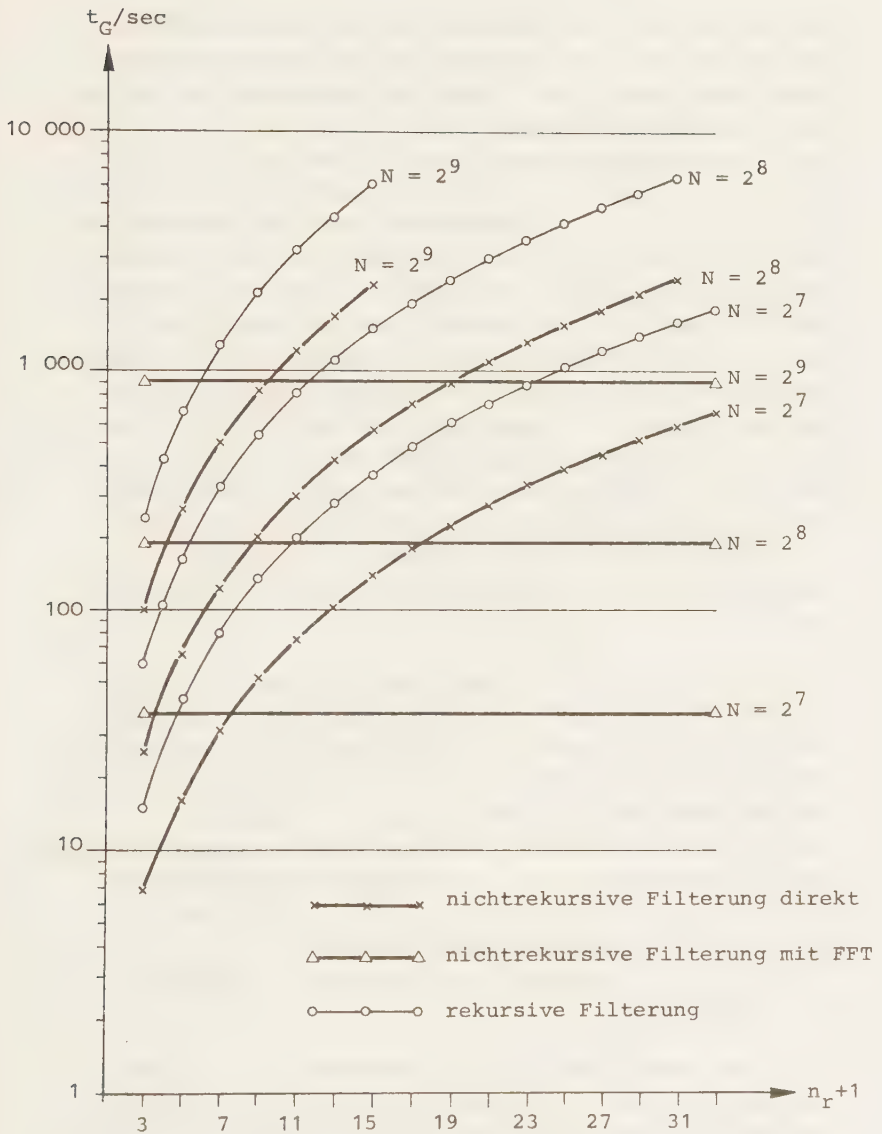


Bild 4.17: Gesamtrechnenzeit t_G in Abhängigkeit von der Filterlänge $n_r + 1$ und der Bildgröße N für $N_B = 2^{13}$

und Impulsantwort erforderlich ist (z.B. 90° -Fan Filter oder Filter mit rotationssymmetrischem Betragsfrequenzgang). Derartige Filter können mit rekursiven Viertelebenfiltern durch Anwendung eines Filters in vier verschiedenen Rekursionsrichtungen (Kaskade oder parallel nach Kap. 4.3.1) realisiert werden. Für diesen Fall ist in dem Diagramm nach Bild 4.18 angegeben, wann die rekursive Filterung gegenüber der nichtrekursiven Filterung Vorteile bringt. Vorausgesetzt wurde dabei, daß Ein- und Ausgangsbild auf $N \times N$, $N=128$ Bildpunkte beschränkt sind, und $N_B=2^{13}$ reelle Datenworte als Hauptspeicherplatz zur Verfügung stehen. Ferner wurde angenommen, daß bei der FFT-Implementierung die Transformationslänge $2N$ beträgt, um Überlappungseffekte an den Rändern (Bild 3.10) zu vermeiden. Bei der direkten Realisierung der nichtrekursiven Filterung wurde die eindimensionale Realisierung (Bild 3.8) und die Reduktion der arithmetischen Operationen durch die vorgegebene Quadrantensymmetrie der Impulsantwort (3.6-6) mit einbezogen.

Läßt sich ein vorgeschriebener Wunschverlauf von Impulsantwort oder Betragsfrequenzgang nach einem vorgegebenen Fehlerkriterium entweder durch ein zweidimensionales nichtrekursives nullphasiges Filter mit quadrantensymmetrischer Impulsantwort und der Länge $(n_{nr} + 1) \cdot (n_{nr} + 1)$ oder durch die vierfache Anwendung eines rekursiven Viertel-ebenenfilters mit Zähler- und Nennermatrizen der Größe $(n_r + 1) \cdot (n_r + 1)$ approximieren, so ist die rekursive Realisierung vorzuziehen, wenn der Punkt $(n_r+1, n_{nr}+1)$ innerhalb des schraffierten Bereichs von Bild 4.18 liegt.

Setzt man voraus, daß für die Approximation eines Wunschverlaufs der Übertragungsfunktion oder der Impulsantwort durch ein nichtrekursives Filter mindestens 10mal mehr Koeffizienten erforderlich sind als bei der Approximation

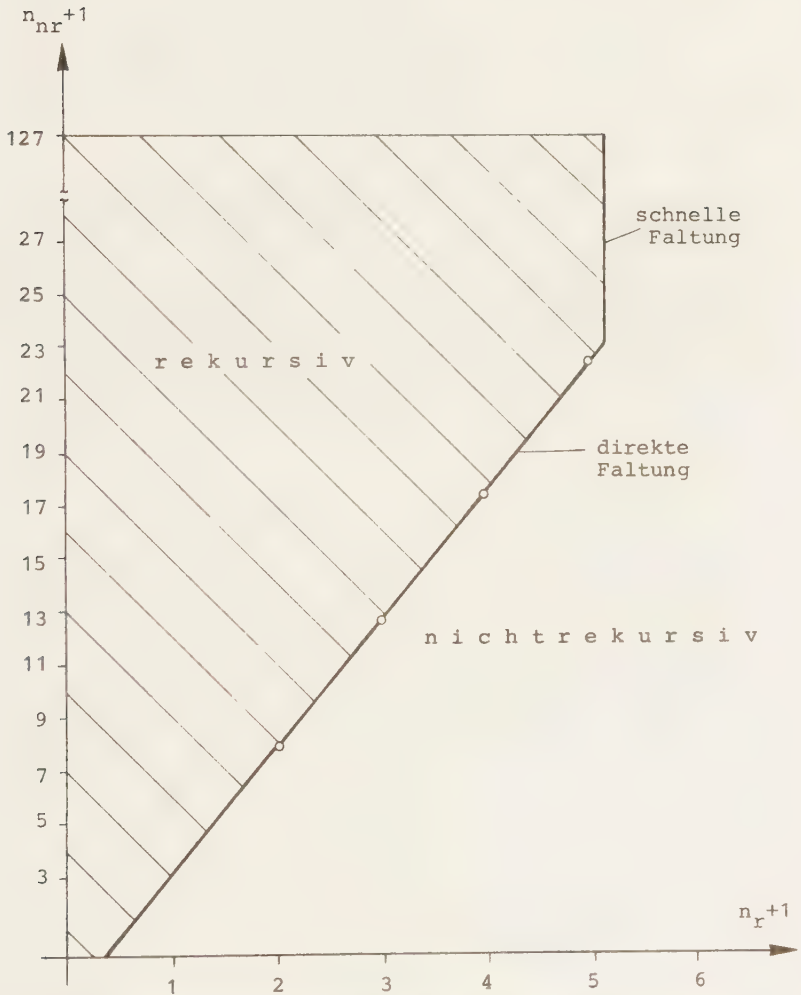


Bild 4.18: Bereiche, in denen die Realisierung durch ein rekursives oder nichtrekursives System Vorteile bringt (für $N = 2^7$, $N_B = 2^{13}$).

durch ein rekursives System [60], gilt für die Filterlänge des nichtrekursiven Systems

$$n_{nr}+1 \geq \text{Max} \left\{ \sqrt{10[2(n_r+1)^2-1]}, n_o \right\},$$

wobei n_o die kleinste ungerade ganze Zahl angibt, die größer oder gleich $\sqrt{10[(2n_r+1)^2-1]}$ ist.

Für diesen Fall ist die rekursive Filterung für $n_r \leq 4$ stets vorteilhafter.

Filter mit speziellen Koeffizienten oder separablem Zähler- und/oder Nennerpolynom sowie Halbebenefilter, die nach Bild 4.6 realisiert werden, sind auch für $n_r > 4$ vergleichbaren nichtrekursiven Filtern überlegen.

5. Entwurf zweidimensionaler rekursiver Filter

Wie im eindimensionalen Fall kann der Entwurf von zweidimensionalen rekursiven digitalen Filtern entweder im

Zeit - bzw. Ortsbereich

oder im

Frequenz - bzw. Ortsfrequenzbereich

erfolgen. Dabei wird ein vorgegebener Sollverlauf einer Impulsantwort oder eines Betragsfrequenzgangs, mit oder ohne Bedingungen an die Phase bzw. Gruppenlaufzeit, entsprechend einer gewählten Fehlernorm im Sinne eines minimalen Fehlers approximiert. Demnach unterscheidet man zwischen Frequenz- und Ortsbereichsentwurfsverfahren. In welchem Bereich der Entwurf erfolgt, hängt vom jeweiligen Anwendungsfall bzw. der Problemstellung ab. Wird ein vorgeschriebener Verlauf der Impulsantwort innerhalb eines endlichen Ortsintervalls gefordert, so ist die Approximation zweckmäßigerweise im Ortsbereich durchzuführen, da ein vorgeschriebener Betragsfrequenzgang allein nicht ausreicht, um eine bestimmte Impulsantwort zu garantieren, und somit müssen Betrags- und Phasenverlauf gleichzeitig approximiert werden, was besonders im zweidimensionalen Fall einen i.a. größeren Aufwand bedeutet. Mit der Z-Transformation $|1|$ können Forderungen im Ortsbereich in äquivalente Forderungen im Frequenzbereich umgerechnet werden und umgekehrt.

5.1 Ortsbereichsentwurfsverfahren

Im Fall des Ortsbereichsentwurfs wird bei den meisten Verfahren der mittlere quadratische Fehler (L_2 -Norm)

$$\Delta = \sum_{k=0}^K \sum_{l=0}^L [h_{oi}(k,l) - h_{oa}(k,l)]^2 \quad (5.1-1)$$

in dem endlichen Intervall $k=0,1,\dots,K, l=0,1,\dots,L$ minimiert. Dabei ist $h_o(m,n)$ die gewünschte Sollimpulsantwort und $h_{o_a}^i(m,n)$ die Impulsantwort des approximierenden Filters.

Zeitbereichsentwurfsverfahren wurden von Shanks [28], Betran [61], Nowrouzian [62], Parker u. Souchon [63], Bordner [64], Mersereau und Shaw [31] veröffentlicht.

In [62] findet sich eine mathematische Beschreibung und in [31] eine ausführliche Diskussion der derzeit bekannten Ortsbereichsentwurfsverfahren. Hier wird deshalb auf eine weitergehende Beschreibung verzichtet.

Der Schwerpunkt dieses Kapitels liegt auf der Beschreibung von neuen Methoden zum Entwurf von rekursiven digitalen Filtern im Frequenzbereich unter Verwendung von eindimensionalen Entwurfsverfahren.

5.2 _Entwurf_im_Frequenzbereich

Der Entwurf im Frequenzbereich ist seit längerer Zeit Hauptgegenstand der Forschung auf dem Gebiet der zweidimensionalen Signalverarbeitung. Die bekannten Verfahren lassen sich unterteilen in Methoden für Viertelebenenfilter [5,50,65-76] und Halbebenenfilter [40,42,46,47,77-79].

Wie in 4.3.2 gezeigt wurde, läßt sich jedes Halbebenenfilter stabilitätserhaltend in ein Viertelebenenfilter mit verschobener Koeffizientenmaske (Bild 4.11) transformieren [33]. Der zu approximierende Wunschfrequenzgang ist dann ebenfalls mit (4.3-6) umzurechnen. Damit läßt sich jedes Halbebenenfilter als Viertelebenenfilter mit spezieller Koeffizientenmaske entwerfen, weshalb es genügt, sich auf Entwurfsverfahren für Viertelebenenfilter zu beschränken.

Bei allen Verfahren muß stets die Stabilität des durch die Approximation erhaltenen rekursiven Filters geprüft (Kap. 4.2), und instabile Filter müssen während oder nach dem Entwurf durch geeignete Maßnahmen [28,31,79] stabilisiert wer-

den. Dies ist jedoch relativ zeitaufwendig.

Da sich Polynome mit mehr als einer Variablen abgesehen von separablen Filtern im allgemeinen nicht in Teilprodukte faktorisieren lassen, hat man bei Entwurfsverfahren für Filter in direkter Form mit gleichem Zähler- und Nennergrad bei voller Ausnutzung aller Freiheitsgrade

$$p = 2(n+1)^2 - 1$$

Parameter, wobei n den Filtergrad angibt. So ergeben sich bei einem Filter 5. Grades bereits 71 Optimierungsparameter. Das bedeutet neben einem hohen Optimierungs- auch einen großen Rechenaufwand für die Realisierung. Im Gegensatz zum 1-D Fall, wo die Aufspaltung in Teilsysteme die Zahl der Freiheitsgrade und den Realisierungsaufwand nicht ändert, reduzieren sich beide Größen für 2-D Filter. Für die Kaskadenschaltung ergeben sich

$$p = 2 \sum_{\lambda=1}^1 (n_{\lambda}+1)^2 - 1$$

freie Parameter, wenn

$$n = \sum_{\lambda=1}^1 n_{\lambda} \quad \text{ist.}$$

Für die Kaskade von Blöcken 1. und 2. Grades ($n_{\lambda} \in \{1, 2\}$) ergeben sich bei $n=5$ 41 freie Optimierungsparameter. Diese Zahl ist noch ausreichend, um eine genügend gute Approximation zu erreichen. Der Rechenaufwand bei der Realisierung reduziert sich proportional zur Zahl der freien Parameter, was ein entscheidender Vorteil ist. Weitere Vorteile gegenüber der direkten Form sind [31, 65]:

- Die Empfindlichkeit der Koeffizienten bezüglich Quantisierungsfehlern ist geringer, was die Stabilität und die Approximation betrifft.
- Es ergibt sich eine einfachere Überprüfung der Stabilität der Teilblöcke 1. und 2. Grades (Kap. 4.2.2.1, [51]).

- Der Datenspeicher, der für die Filterung benötigt wird, kann kleiner gehalten werden.

Aus diesen Gründen wird häufig eine Kaskadierung oder Parallelschaltung von zweidimensionalen Blöcken 1. und 2. Grades vorgezogen [50,65-67,74-79].

5.3 _Neue Verfahren zum Entwurf zweidimensionaler rekursiver stabiler Filter

Bei allen in diesem Abschnitt beschriebenen Verfahren wird von einem eindimensionalen rekursiven digitalen Filter vom Grad n mit der Übertragungsfunktion

$$H(z) = \frac{\sum_{\mu=0}^n b_{\mu} z^{\mu}}{\sum_{\mu=0}^n c_{\mu} z^{\mu}} \quad (5.3-1)$$

und $c_n=1$, $z=e^{j\Omega}$, $\Omega=2\pi f/f_A$, b_{μ} , c_{μ} reell, f_A =Abtastfrequenz ausgegangen.

Eine derartige eindimensionale Übertragungsfunktion läßt sich stets in eine Parallelförm (4.kanonische Form)

$$H(z) = b_0 + \sum_{\lambda=1}^L \tilde{H}_{\lambda}(z) \quad (5.3-2)$$

überführen, wobei $\tilde{H}_{\lambda}(z)$ die Übertragungsfunktion eines Blocks 1. (bei reellen Polen) oder 2. Grades (bei komplexen Polpaaren) bedeutet. Eine alternative Darstellung ist die Kaskade von Blöcken 1. und 2. Grades (3.kanonische Form) mit den Übertragungsfunktionen $H_{\lambda}(z)$ und

$$H(z) = \prod_{\lambda=1}^l H_{\lambda}(z) \quad . \quad (5.3-3)$$

Dabei gilt $l=n-n_2/2$, wobei $n_2/2$ die Zahl der komplexen Polpaare angibt $|1|$.

5.3.1 Entwurfsverfahren 1

5.3.1.1 Transformation eines Blocks 2.Grades

Betrachtet sei nun einer der Teilblöcke 2.Grades mit der Übertragungsfunktion

$$H_1(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^2 + c_1 z + c_0} \quad . \quad (5.3-4)$$

Mit Hilfe der bilinearen Transformation $|1|$

$$z = \frac{1+w}{1-w} \quad (5.3-5)$$

ergibt sich der entsprechende Block 2.Grades im w -Bereich

$$G_1(w) = \frac{a_2 w^2 + a_1 w + a_0}{d_2 w^2 + d_1 w + d_0} = \frac{A_1(w)}{D_1(w)} \quad , \quad (5.3-6)$$

wobei $a_2 = b_2 - b_1 + b_0$

$$a_1 = 2(b_2 - b_0)$$

$$a_0 = b_2 + b_1 + b_0$$

(5.3-7)

$$d_2 = 1 - c_1 + c_0$$

$$d_1 = 2(1 - c_0)$$

$$d_0 = 1 + c_1 + c_0 \quad \text{ist.}$$

Den Übergang in den z -Bereich erreichen wir, wenn wir im allgemeinsten Fall die Frequenzvariable w durch Funktionen

von w_1 und w_2 ersetzen. Um eine i.a. unerwünschte Graderweiterung zu vermeiden, darf sich der Zähler und Nennergrad bzgl. w_1, w_2 nach der Transformation nicht ändern.

Diese Bedingung wird auch erfüllt, wenn w bzw. w^2 getrennt für Zähler und Nenner von (5.3-6) durch zweidimensionale Polynome 1. und 2. Grades substituiert werden. Hier wird der Zähler von (5.3-6) mit

$$\begin{aligned} w^2 &\rightarrow P(w_1, w_2) = \sum_{\mu=0}^2 \sum_{\nu=0}^2 p_{\mu\nu} w_1^\mu w_2^\nu \\ w &\rightarrow Q(w_1, w_2) = \sum_{\mu=0}^1 \sum_{\nu=0}^1 q_{\mu\nu} w_1^\mu w_2^\nu \end{aligned} \quad (5.3-8a)$$

und der Nenner mit

$$\begin{aligned} w^2 &\rightarrow R(w_1, w_2) = \sum_{\mu=0}^2 \sum_{\nu=0}^2 r_{\mu\nu} w_1^\mu w_2^\nu \\ w &\rightarrow T(w_1, w_2) = \sum_{\mu=0}^1 \sum_{\nu=0}^1 t_{\mu\nu} w_1^\mu w_2^\nu \end{aligned} \quad (5.3-8b)$$

transformiert. Die Transformationskoeffizienten $p_{\mu\nu}$, $q_{\mu\nu}$, $r_{\mu\nu}$, $t_{\mu\nu}$ sind so zu bestimmen, daß nicht nur die gestellten Forderungen an das Frequenzverhalten erfüllt werden, sondern auch die Stabilität des Filters erhalten bleibt. Dies führt schließlich zu der zweidimensionalen Funktion

$$G_2(w_1, w_2) = \frac{\underline{w}_1^T [\beta_{\mu\nu}] \underline{w}_2}{\underline{w}_1^T [\zeta_{\mu\nu}] \underline{w}_2} \quad (5.3-9)$$

$$\text{mit } \underline{w}_\rho^T = [1 \ w_\rho \ w_\rho^2]$$

$$\begin{aligned} \mu, \nu &\in \{0, 1, 2\} \\ \rho &\in \{1, 2\} \end{aligned} .$$

Die Koeffizienten $\beta_{\mu\nu}$ und $\zeta_{\mu\nu}$ des zweidimensionalen Blocks 2. Grades $G_2(w_1, w_2)$ hängen dabei sowohl von den Koeffizienten der eindimensionalen Funktion $G_1(w)$ (5.3-6) als auch von den Transformationsparametern (5.3-8a,b) ab [siehe z.B. (5.3-13)].

Durch spezielle Forderungen läßt sich die Zahl der Transformationsparameter reduzieren.

Häufig ist eine Symmetrie der Übertragungsfunktion bezüglich der frequenzabhängigen Variablen z_1, z_2 bzw. w_1, w_2 gefordert. Gilt

$$G_2(w_1, w_2) = G_2(w_2, w_1) , \quad (5.3-10)$$

so folgt daraus

$$\begin{aligned} \beta_{\mu\nu} &= \beta_{\nu\mu} \\ \zeta_{\mu\nu} &= \zeta_{\nu\mu} . \end{aligned}$$

Fordert man außerdem für $(w_1, 0)$ den gleichen Verlauf der Übertragungsfunktion wie im eindimensionalen Fall, d.h.

$$G_2(w_1, 0) \stackrel{!}{=} G_1(w_1) , \quad (5.3-11)$$

so ergibt sich

$$G_2(w_1, w_2) = \frac{\underline{w}_1^T \cdot \begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & \beta_{11} & \beta_{12} \\ a_2 & \beta_{12} & \beta_{22} \end{bmatrix} \cdot \underline{w}_2}{\underline{w}_1^T \cdot \begin{bmatrix} d_0 & d_1 & d_2 \\ d_1 & \zeta_{11} & \zeta_{12} \\ d_2 & \zeta_{12} & \zeta_{22} \end{bmatrix} \cdot \underline{w}_2} = \frac{A_2(w_1, w_2)}{D_2(w_1, w_2)} , \quad (5.3-12)$$

wobei sich die a_v und $d_v, v \in \{0, 1, 2\}$ nach (5.3-7) berechnen und

$$\begin{aligned}
\beta_{11} &= a_2 p_{11} + a_1 q_{11} \\
\beta_{12} &= a_2 p_{21} = a_2 p_{12} \\
\beta_{22} &= a_2 p_{22} \\
\zeta_{11} &= d_2 r_{11} + d_1 t_{11} \\
\zeta_{12} &= d_2 r_{21} = d_2 r_{12} \\
\zeta_{22} &= d_2 r_{22}
\end{aligned}
\tag{5.3-13}$$

ist.

Wegen Gl. (5.3-10) gilt dann auch $G_2(0, w_2) = G_1(w_2)$.

Durch Anwendung der bilinearen Rücktransformation für beide Frequenzvariable w_1, w_2 in (5.3-12) mit

$$w_\rho = \frac{z_\rho^{-1}}{z_\rho + 1} \quad \rho \in \{1, 2\} \tag{5.3-14}$$

ergeben sich die Koeffizienten des zweidimensionalen digitalen Filterblocks 2. Grades

$$H_2(z_1, z_2) = \frac{\underline{z}_1^T \underline{B} \underline{z}_2}{\underline{z}_1^T \underline{C} \underline{z}_2} = \frac{B(z_1, z_2)}{C(z_1, z_2)} \tag{5.3-15}$$

mit $\underline{z}_\rho = [1, z_\rho^{-1}, z_\rho^{-2}]^T$, $\rho \in \{1, 2\}$.

Die Koeffizientenmatrizen $\underline{B}, \underline{C}$ sind als Folge von Forderung (5.3-10) symmetrisch.

D.h. $\underline{B} = [b_{\mu\nu}]$, $b_{\mu\nu} = b_{\nu\mu}$

$\underline{C} = [c_{\mu\nu}]$, $c_{\mu\nu} = c_{\nu\mu}$

mit $\mu, \nu \in \{0, 1, 2\}$.

$$\begin{aligned}
\text{Mit } \underline{b}^{(2)} &= [b_{00} \ b_{01} \ b_{02} \ b_{11} \ b_{12} \ b_{22}]^T \\
\underline{c}^{(2)} &= [c_{00} \ c_{01} \ c_{02} \ c_{11} \ c_{12} \ c_{22}]^T \\
\underline{\beta} &= [\beta_{11} \ \beta_{12} \ \beta_{22}]^T \\
\underline{\zeta} &= [\zeta_{11} \ \zeta_{12} \ \zeta_{22}]^T \\
\underline{b}^{(1)} &= [b_0 \ b_1 \ b_2]^T \\
\underline{c}^{(1)} &= [c_0 \ c_1 \ 1]^T
\end{aligned} \tag{5.3-16}$$

berechnen sich die Koeffizientenvektoren $\underline{b}^{(2)}$, $\underline{c}^{(2)}$ zu

$$\underline{b}^{(2)} = \underline{T}_1 \underline{\beta} + \underline{T}_2 \underline{b}^{(1)} \tag{5.3-17a}$$

bzw.

$$\underline{c}^{(2)} = \underline{T}_1 \underline{\zeta} + \underline{T}_2 \underline{c}^{(1)}, \tag{5.3-17b}$$

wobei die Vektoren $\underline{b}^{(1)}$, $\underline{c}^{(1)}$ durch (5.3-4) und $\underline{\beta}$, $\underline{\zeta}$ durch (5.3-13) mit (5.3-16) gegeben sind.

Die Transformationsmatrizen $\underline{T}_1, \underline{T}_2$ berechnen sich zu

$$\underline{T}_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & -2 \\ -1 & 0 & 1 \\ 0 & 0 & 4 \\ 0 & 2 & -2 \\ 1 & -2 & 1 \end{bmatrix} \tag{5.3-18a}$$

und

$$\underline{T}_2 = \begin{bmatrix} -1 & -1 & 7 \\ -2 & 2 & 6 \\ 3 & -1 & 3 \\ -4 & 12 & -4 \\ 6 & 2 & -2 \\ 7 & -1 & -1 \end{bmatrix} \tag{5.3-18b}$$

Betrachtet sei nun der Fall, daß

$$b_2 = b_0 = b_1/2$$

gilt. Dies bedeutet, daß der eindimensionale Filterblock $H_1(z)$ nach (5.3-4) bei $z = -1$ stets eine doppelte Nullstelle besitzt, was z.B. bei Filtern mit Potenzverhalten bzw. Tschebyscheff-Verhalten im Durchlaßbereich stets gegeben ist.

Nach (5.3-7) ist dann $a_2 = a_1 = 0$, d.h. in diesem Fall wird $\beta_{11} = \beta_{12} = \beta_{22} = a_2 = a_1 = 0$ und $a_0 = 4 \cdot b_0$.

Es ergibt sich

$$G_2(w_1, w_2) = \frac{4 \cdot b_0}{\underline{w}_1^T \begin{bmatrix} d_0 & d_1 & d_2 \\ d_1 & \zeta_{11} & \zeta_{12} \\ d_2 & \zeta_{12} & \zeta_{22} \end{bmatrix} \underline{w}_2} \quad (5.3-19)$$

Für diesen Sonderfall wird das Zählerpolynom $B(z_1, z_2)$ von $H_2(z_1, z_2)$ in (5.3-15) separabel.

Dann gilt:

$$B(z_1, z_2) = 4 \cdot b_0 (z_1^{-1} - 1)^2 \cdot (z_2^{-1} - 1)^2. \quad (5.3-20)$$

Für den Fall, daß $b_0 = b_2$ ist, wodurch die Nullstellen des Zählers im 1-D Fall auf dem Einheitskreis liegen, ergibt sich für $\beta_{11} = \beta_{12} = 0$ eine oktandensymmetrische Zählerkoeffizientenmatrix

$$\underline{B} = \begin{bmatrix} b_{00} & b_{10} & b_{00} \\ b_{10} & b_{11} & b_{10} \\ b_{00} & b_{10} & b_{00} \end{bmatrix}. \quad (5.3-21)$$

Diese Sonderfälle sind insbesondere beim Entwurf von Filtern mit zirkulärsymmetrischem Betragsfrequenzgang interessant, da sie zum einen die Zahl der arithmetischen Operationen bei der Realisierung reduzieren und zum anderen

schon für eine gewisse Symmetrie des Betragsfrequenzgangs sorgen.

Zusammenfassend läßt sich feststellen:

- Pro Block 2.Grades ergeben sich je 3 freie Entwurfparameter für Zähler und Nenner, die so zu wählen sind, daß die gestellten Forderungen an das Frequenzverhalten erfüllt werden und das System stabil bleibt.
Rechnet man die 5 Parameter b_2, b_1, b_0, c_1, c_0 des 1-D Systems nach (5.3-4) hinzu, so ergeben sich 11 Parameter für die Optimierung, also genausoviel wie für eine direkte Optimierung eines Blocks 2.Grades mit symmetrischen Koeffizientenmatrizen. Ein Vorteil dieser Vorgehensweise liegt darin, daß der Entwurf- bzw. die Optimierung zum Teil mit Verfahren für den 1-D Entwurf durchgeführt werden kann.
- Für den Fall einer doppelten Nullstelle bei $z=-1$ liegt das Zählerpolynom als separables 2-D Polynom fest (5.3-20), d.h. die Zahl der Entwurfparameter reduziert sich um 3.
- Für den Fall, daß die Nullstellen des eindimensionalen Filters auf dem Einheitskreis liegen ($b_0=b_2$), kann durch Beschränkung auf einen freien Parameter für den Zähler erreicht werden, daß die Zählerkoeffizientenmatrix oktandensymmetrisch wird.

5.3.1.2 Untersuchung der Stabilität

Zu untersuchen ist nun, welche Bedingungen die Transformationsparameter $\zeta_{11}, \zeta_{12}, \zeta_{22}$ erfüllen müssen, damit der durch die Transformation erhaltene zweidimensionale Filterblock $G_2(w_1, w_2)$ bzw. $H_2(z_1, z_2)$ (5.3-9, 5.3-15) stabil bleibt, sofern das eindimensionale Filter stabil ist.

Um bei der Rücktransformation vom (w_1, w_2) - in den (z_1, z_2) -Bereich mit der bilinearen Transformation (5.3-14) unwesentliche Singularitäten 2.Art bei $(z_1, z_2)=(-1, -1)$ zu vermeiden $|32|$, muß

$$\zeta_{22} \neq 0 \text{ sein.}$$

Das resultierende Filter ist genau dann stabil, wenn das Nennerpolynom $D_2(w_1, w_2)$ nach (5.3-12) ein stabiles zweidimensionales Polynom darstellt. Dies ist der Fall, wenn

1. $D_2(w_1, 1) \neq 0$ für $\operatorname{Re}\{w_1\} \geq 0$
 2. $D_2(jv_1, w_2) \neq 0$ für $\operatorname{Re}\{w_2\} \geq 0$, $v_1 \in \mathbb{R}$
- (5.3-22)

gilt |81|.

Eine Ableitung der für die Sicherstellung der Stabilität zu erfüllenden Bedingungen für die Transformationsparameter ζ_{11}, ζ_{22} ist in dieser Form nicht mehr allgemein möglich. Deshalb wird das in 5.3.1.1 beschriebene Verfahren modifiziert.

5.3.1.3 Transformation eines Blocks 2. Grades mit komplexen Transformationsparametern

Wir betrachten nun ein 1-D Teilfilter 2. Grades mit konjugiert komplexen Null- und Polstellen mit der Teilübertragungsfunktion

$$H_1(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^2 + c_1 z + c_0} = b_2 \cdot \frac{z - z_0}{z - z_\infty} \cdot \frac{z - z_0^*}{z - z_\infty^*} \quad (5.3-23)$$

Mit der bilinearen Transformation (5.3-5) ergibt sich das Nennerpolynom von (5.3-6) zu

$$D_1(w) = d_0 \left(1 + w \frac{1 + z_\infty}{1 - z_\infty}\right) \left(1 + w \frac{1 + z_\infty^*}{1 - z_\infty^*}\right) \quad (5.3-24)$$

Als Sonderfall von (5.3-8b) wird w jetzt getrennt für jeden Teilfaktor mit komplexen Parametern transformiert. Unter Berücksichtigung der Bedingungen (5.3-10), (5.3-11) ergibt sich aus (5.3-8b)

$$t_{00}=0$$

$$t_{01}=t_{10}=1.$$

$$\text{Mit } t_{11}=(\alpha+j\beta) \frac{1-z_{\infty}}{1+z_{\infty}} \quad (5.3-25)$$

erhält man dann das Nennerpolynom $D_2(w_1, w_2)$ in (5.3-12) zu

$$\begin{aligned} D_2(w_1, w_2) &= d_0 \{ 1 + [t_{11} w_1 w_2 + t_{10} (w_1 + w_2) + t_{00}] \} \\ &\quad \cdot \{ 1 + [t_{11}^* w_1 w_2 + t_{10}^* (w_1 + w_2) + t_{00}^*] \} \\ &= d_0 [1 + (\alpha + j\beta) w_1 w_2 + \frac{1+z_{\infty}}{1-z_{\infty}} (w_1 + w_2)] \\ &\quad \cdot [1 + (\alpha - j\beta) w_1 w_2 + \frac{1+z_{\infty}^*}{1+z_{\infty}^*} (w_1 + w_2)]. \end{aligned} \quad (5.3-26)$$

$$\begin{aligned} \text{Mit } d_0 &= (1-z_{\infty})(1-z_{\infty}^*) = 1+c_1+c_0 \\ d_1 &= (1-z_{\infty})(1+z_{\infty}^*) + (1-z_{\infty}^*)(1+z_{\infty}) = \\ &= 2(1-c_0) \\ d_2 &= (1+z_{\infty})(1+z_{\infty}^*) = 1-c_1+c_0 \\ d_3 &= 2 \operatorname{Im}\{z_{\infty}\} = 2 \sqrt{c_0 - c_1^2/4} \end{aligned} \quad (5.3-27)$$

ergibt sich aus (5.3-26)

$$D_2(w_1, w_2) = \underline{w}_1^T \cdot \begin{bmatrix} d_0 & d_1 & d_2 \\ d_1 & 2(\alpha d_0 + d_2) & \alpha d_1 + 2d_3\beta \\ d_2 & \alpha d_1 + 2d_3\beta & d_0(\alpha^2 + \beta^2) \end{bmatrix} \cdot \underline{w}_2 \quad (5.3-28)$$

wobei $\underline{w}_\rho = [1 \ w_\rho \ w_\rho^2]^T$, $\rho \in \{1, 2\}$ ist.

Ein Vergleich mit (5.3-12) zeigt, daß jetzt

$$\begin{aligned} \zeta_{11} &= 2(\alpha d_0 + d_2) \\ \zeta_{12} &= \alpha d_1 + 2d_3\beta \\ \zeta_{22} &= d_0(\alpha^2 + \beta^2) \end{aligned} \quad (5.3-29)$$

gilt.

Die Zahl der für den Nenner frei wählbaren Transformationsparameter hat sich damit auf 2 reduziert. Der Vorteil liegt darin, daß für die Stabilitätsbedingungen angegeben werden können. Hierbei sind die 2 freien Parameter für das 1-D Prototypfilter nicht hinzugerechnet, sie liegen nach dem 1-D Entwurf fest.

5.3.1.4 Stabilitätsbedingungen für die Transformationsparameter

Das Nennerpolynom $D_2(w_1, w_2)$ nach (5.3-26) ist offenbar genau dann ein stabiles Polynom, d.h. die Bed. (5.3-22) werden erfüllt, wenn das Teilpolynom mit komplexen Koeffizienten

$$N(w_1, w_2) = 1 + (\alpha + j\beta) \cdot w_1 \cdot w_2 + \frac{1+z_\infty}{1-z_\infty} \cdot (w_1 + w_2) \quad (5.3-30)$$

stabil ist. Daraus sind die Bedingungen für α und β herzuleiten.

Es gilt:

$$\frac{1+z_\infty}{1-z_\infty} = \frac{1}{d_0} \left(\frac{d_1}{2} + jd_3 \right), \quad (5.3-31)$$

wobei d_0, d_1, d_3 durch Gl. (5.3-27) gegeben sind.

$$\begin{aligned} \text{Damit } N_2(w_1, w_2) &= N(w_1, w_2) \cdot d_0 = \\ &= d_0 + (\alpha + j\beta) \cdot d_0 \cdot w_1 \cdot w_2 + \left(\frac{d_1}{2} + jd_3 \right) (w_1 + w_2) \end{aligned} \quad (5.3-32)$$

ein stabiles Polynom ist, müssen die Bedingungen analog zu (5.3-22)

$$1. \quad N_2(w_1, 1) \neq 0 \quad \text{für } \operatorname{Re}\{w_1\} \geq 0 \quad (5.3-33a)$$

und

$$2. \quad N_2(jv_1, w_2) \neq 0 \quad \text{für } \operatorname{Re}\{w_2\} \geq 0, v_1 \in \mathbb{R} \quad (5.3-33b)$$

eingehalten werden $|81|$.

Durch Berechnung der Nullstellen der Polynome $N_2(w_1, 1)$ und $N_2(jv_1, w_2)$ ergibt sich zunächst mit der Bedingung (5.3-33a)

$$(d_0 + \frac{d_1}{2}) \cdot (\alpha d_0 + \frac{d_1}{2}) + d_3(\beta d_0 + d_3) > 0 \quad . \quad (5.3-34)$$

Aus der 2. Bedingung (5.3-33b) folgt

$$(\beta d_3 + \frac{d_1}{2}\alpha) \cdot v_1^2 - d_0 \beta v_1 + \frac{d_1}{2} > 0 \quad . \quad (5.3-35)$$

Damit die Ungleichung (5.3-35) für alle $v_1 \in \mathbb{R}$ erfüllt ist, darf die entsprechende quadratische Gleichung für v_1 keine reelle Lösung besitzen. Dies erfordert, daß

$$2 d_1 (\beta d_3 + \frac{d_1}{2}\alpha) - d_0^2 \beta^2 > 0 \quad (5.3-36)$$

ist.

Aus der Ungleichung (5.3-36) läßt sich die obere und untere Schranke für den Transformationsparameter β berechnen.

Es ergibt sich

$$\beta_{\min} < \beta < \beta_{\max} \quad (5.3-37a)$$

$$\text{mit } \beta_{\min} = \frac{d_1 d_3}{d_0^2} - \frac{d_1}{d_0} \sqrt{\left(\frac{d_3}{d_0}\right)^2 + \alpha} \quad (5.3-37b)$$

$$\beta_{\max} = \frac{d_1 d_3}{d_0^2} + \frac{d_1}{d_0} \sqrt{\left(\frac{d_3}{d_0}\right)^2 + \alpha} \quad (5.3-37c)$$

da nach (5.3-27) $d_0, d_1 > 0$, $d_3 \geq 0$ gilt, falls das eindimensionale Filter stabil ist.

Außerdem muß

$$\alpha > - \left(\frac{d_3}{d_0}\right)^2 \quad (5.3-38)$$

sein, da sonst die Ungleichung (5.3-36) verletzt wird.

Mit

$$\Delta\beta = \beta_{\max} - \beta_{\min} = \frac{2d_1}{d_0} \sqrt{\left(\frac{d_3}{d_0}\right)^2 + \alpha}, \quad (5.3-39)$$

$$\alpha = -\left(\frac{d_3}{d_0}\right)^2 + \varepsilon + x^2, \quad x \in \mathbb{R}, \quad (5.3-40)$$

$$\beta = \beta_{\min} + \varepsilon + (\Delta\beta - 2\varepsilon) \sin^2 \phi \quad (5.3-41)$$

$$0 < \varepsilon \ll 1 \quad (5.3-42)$$

werden die Ungleichungen (5.3-34) bis (5.3-38) stets erfüllt (siehe Anhang 2).

Aus (5.3-39) bis (5.3-42) ergibt sich schließlich

$$\alpha = -\left(\frac{d_3}{d_0}\right)^2 + x^2 + \varepsilon, \quad (5.3-43)$$

$$\beta = \frac{d_1 d_3}{d_0^2} + \left(\frac{d_1}{d_0} \sqrt{x^2 + \varepsilon} - \varepsilon\right) (2 \sin^2 \phi - 1). \quad (5.3-44)$$

Hiermit sind die eingeschränkten Transformationsparameter α, β in zwei nicht eingeschränkte reelle Entwurfparameter x, ϕ übergeführt worden. Die Konstante ε sichert einen Minimalabstand der Pole von der imaginären Achse (im w -Bereich) bzw. vom Einheitskreis (im z -Bereich) und damit die Stabilität des transformierten Systems. Zweckmäßigerweise ist ε zwischen $10^{-3} \dots 10^{-5}$ zu wählen.

Zu klären ist nun noch, wie Blöcke 1. Grades in (5.3-2) bzw. (5.3-3) transformiert werden können. Dazu hat man zwei Möglichkeiten. Die erste Möglichkeit ist, für den Block 1. Grades Zähler und Nenner mit z zu multiplizieren.

Das ergibt

$$H(z) = \frac{\hat{b}_1 z + \hat{b}_0}{z + \hat{c}_0} = \frac{b_2 z^2 + b_1 z}{z^2 + c_1 z}$$

mit $b_2 = \hat{b}_1$, $b_1 = \hat{b}_0$, $c_1 = \hat{c}_0$.

Damit erhält man einen Block 2.Grades mit den speziellen Koeffizienten $b_0=0$, $c_0=0$. In diesem Fall können alle Beziehungen für den Block 2.Grades übernommen werden, wobei jedoch $d_3=0$ zu setzen ist. Allerdings ergibt sich so als transformiertes 2-D Filter ein Block 2.Grades, was eine unerwünschte Graderhöhung bedeutet. Vorteilhaft ist, daß dadurch ebenfalls 5 freie Entwurfsparameter pro Block zur Verfügung stehen, genau wie beim Block 2.Grades.

Die zweite Möglichkeit ist, den Block 1.Grades getrennt zu transformieren.

Analog zu (5.3-12) ergibt sich

$$G_2(w_1, w_2) = \frac{[1 \ w_1] \begin{bmatrix} a_0 & a_1 \\ a_1 & \beta_{11} \end{bmatrix} \begin{bmatrix} 1 \\ w_2 \end{bmatrix}}{[1 \ w_2] \begin{bmatrix} d_0 & d_1 \\ d_1 & \alpha d_0 \end{bmatrix} \begin{bmatrix} 1 \\ w_2 \end{bmatrix}}, \quad (5.3-45)$$

wobei

$$\begin{aligned} a_0 &= b_1 + b_0 \\ a_1 &= b_1 - b_0 \\ d_0 &= 1 + c_0 \\ d_1 &= 1 - c_0 \end{aligned} \quad (5.3-46)$$

und

$$\alpha = x^2 + \varepsilon, \quad 0 < \varepsilon \ll 1, \quad x \text{ reell} \quad (5.3-47)$$

ist.

Falls α nach (5.3-47) gewählt wird, bleibt das zweidimensionale Filter stets stabil.

Mit den reellen freien Parametern x, β_{11} erhält man pro Block 1. Grades 2 freie Entwurfparameter.

Der zugehörige digitale Block 1. Grades berechnet sich mit der bilinearen Transformation (5.3-14) zu

$$H(z_1, z_2) = \frac{z_1^T \underline{B} z_2}{z_1^T \underline{C} z_2},$$

wobei $\underline{z}_\rho = [1 \ z_\rho^{-1}]^T$ mit $\rho \in \{1, 2\}$ gilt, und die Koeffizientenmatrizen

$$\underline{B} = \begin{bmatrix} \beta_{11} + 3b_1 + b_0 & b_1 + b_0 - \beta_{11} \\ b_1 + b_0 - \beta_{11} & \beta_{11} - b_1 + 3b_0 \end{bmatrix}, \quad (5.3-48)$$

$$\underline{C} = \begin{bmatrix} \alpha(1+c_0) + 3c_0 & (1-\alpha)(1+c_0) \\ (1-\alpha)(1+c_0) & \alpha(1+c_0) + 3c_0 - 1 \end{bmatrix}$$

sind.

Dieses Verfahren eignet sich besonders gut für den Entwurf von Filtern mit zirkulärsymmetrischem Betragsfrequenzgang. Durch den Entwurf des 1-D Prototypfilters liegen die 5 Parameter b_2, b_1, b_0, c_1, c_0 und damit $a_0, a_1, a_2, d_0, d_1, d_2, d_3$ für einen Block 2. Grades fest. Die 5 verbleibenden freien Optimierungsparameter $\beta_{11}, \beta_{12}, \beta_{22}$ für den Zähler (5.3-12) und die abgeleiteten freien Parameter x und ϕ nach (5.3-43) und (5.3-44) für den Nenner unterliegen keinen Einschränkungen und können für jeden Block 2. Grades unterschiedlich gewählt werden. Gilt für die Zählerkoeffizienten des 1-D Blocks 2. Grades $b_0 = b_2 = b_1/2$, so ist das Zählerpolynom nach (5.3-20) bestimmt, und es brauchen nur die zwei Parameter für den Nenner optimiert werden. Die verbleibenden Parameter werden so optimiert, daß sich eine möglichst gute Zirkular-

symmetrie ergibt, ohne den vorher bestimmten Frequenzgang auf den Hauptachsen $(\Omega_1, 0), (0, \Omega_2)$ zu verändern und im Gegensatz zu bisher bekannten Verfahren [50,72] mit garantierter Stabilität für das zweidimensionale Filter. Ein Stabilitätstest oder eine aufwendige Stabilisierung entfällt deshalb.

Zur Unterscheidung mit weiteren noch zu beschreibenden Verfahren wird dieses Verfahren fortan als 'Entwurfsverfahren 1' bezeichnet.

5.3.1.5 Beispiele

Für das vorstehende Entwurfsverfahren wurde an einer PDP11/45 ein FORTRAN-Programm erstellt. Für den Entwurf des eindimensionalen Prototypfilters wurde das Programm DOREDI [82], für die Optimierung der freien Transformationsparameter ein Simplex-Verfahren [58] verwendet. Diese Programme können ohne Einschränkung durch andere Programme ersetzt werden (z.B. [57]). Der Entwurf des 1-D-Prototypfilters erfolgt so, daß ein vorgegebenes Toleranzschema eingehalten wird. Die verbleibenden Transformationsparameter für den 2-D Entwurf werden so optimiert, daß der Approximationsfehler Δ nach der

Maximum (L_∞) -Norm

$$\Delta_\infty = \text{Max } W(\Omega_r) \cdot |S(\Omega_r) - H(\Omega_{1m}, \Omega_{2n})|$$

oder der L_p -Norm

$$\Delta_p = \left\{ \sum_{m=0}^N \sum_{n=0}^N W(\Omega_r) \cdot |S(\Omega_r) - H(\Omega_{1m}, \Omega_{2n})|^p \right\}^{1/p}$$

minimal wird, wobei $m, n = 0, 1, \dots, N$, $r = \sqrt{m^2 + n^2}$ und

$$\Omega_r = r \cdot \frac{\pi}{N}, \quad \Omega_{1m} = m \cdot \frac{\pi}{N}, \quad \Omega_{2n} = n \cdot \frac{\pi}{N} \text{ ist.}$$

Dabei gibt

$S(\Omega_r)$ die Sollfunktion an der Stelle Ω_r ,

$W(\Omega_r)$ eine vorgebbare Gewichtsfunktion an der Stelle Ω_r ,

und $H(\Omega_{1m}, \Omega_{2n})$ den Frequenzgang des 2-D Filters im Punkt $(\Omega_{1m}, \Omega_{2n})$ an.

Speziell für den Entwurf von Filtern mit rotationssymmetrischem Betragsfrequenzgang ist es zweckmäßig, die Sollfunktion nicht auf äquidistanten Gitterpunkten eines kartesischen Koordinatensystems, sondern für Polarkoordinaten vorzugeben. Wegen der geforderten Symmetrie bzgl. z_1 und z_2 (5.3-10) kann die Fehlerberechnung auf den 2. und 3. Oktanten der (Ω_1, Ω_2) -Ebene beschränkt ($z_v = e^{j\Omega_v}$, $\Omega_v = 2\pi f_v / f_{Av}$, $v \in \{1, 2\}$) werden.

Für die nachfolgenden Beispiele wird die L_2 -Norm als Fehlerkriterium gewählt.

5.3.1.5.1 Beispiel 1: Potenztiefpas

Als erstes Beispiel wird ein Potenztiefpas 2. Grades entworfen. In Bild 5.1 sind der Betragsfrequenzgang und die Übertragungsfunktion des eindimensionalen Filters angegeben. Dieses Filter hat eine doppelte Nullstelle bei $z = -1$. Deshalb liegt das Zählerpolynom des zweidimensionalen Filters nach (5.3-20) fest und nur die zwei freien Parameter x und ϕ in (5.3-43), (5.3-44) müssen optimiert werden. Die Nennerkoeffizientenmatrix des zweidimensionalen rekursiven Tiefpasses berechnet sich dann aus (5.3-17b) mit (5.3-18a,b), (5.3-27), (5.3-29), (5.3-43) und (5.3-44). Bild 5.2 zeigt den Betragsfrequenzgang des zweidimensionalen Filters nach der Optimierung der zwei freien Parameter, Bild 5.3a) die zugehörigen Höhenlinien. Optimiert wurde das Filter nach der mittleren quadratischen Abweichung des Betragsfrequenzgangs vom Sollverlauf (L_2 -Norm). Obwohl mit einem stabilen zweidimensionalen Filter, welches nur in einer Rekursionsrichtung angewendet wird, mit Ausnahme von separablen Filtern keine exakte Quadrantensymmetrie erreicht werden kann [29], ist der zirkularsymmetri-

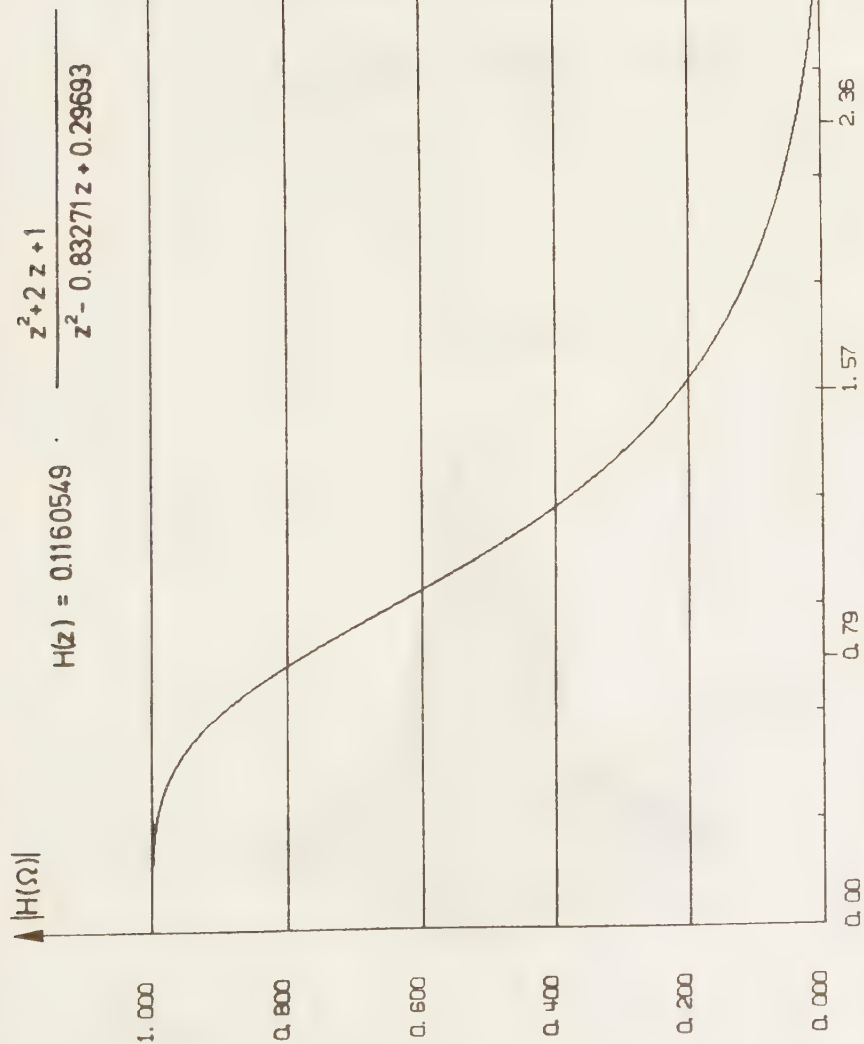


Bild 5.1: Betragsfrequenzgang des 1-D Prototyp-Tiefpasses von Beispiel 1

KONSTANTE = 0.1160549E+00

$$\begin{array}{ccc} 0.4000000E+01 & 0.8000000E+01 & 0.4000000E+01 \\ 0.8000000E+01 & 0.1600000E+02 & 0.8000000E+01 \\ 0.4000000E+01 & 0.8000000E+01 & 0.4000000E+01 \end{array} \left. \vphantom{\begin{array}{ccc} 0.4000000E+01 \\ 0.8000000E+01 \\ 0.4000000E+01 \end{array}} \right\} \underline{B} = [b_{\mu\nu}]$$

$$\begin{array}{ccc} 0.7967131E+00 & -0.2937448E+01 & 0.6891613E+01 \\ -0.2937448E+01 & 0.1428214E+02 & -0.2466805E+02 \\ 0.6891613E+01 & -0.2466805E+02 & 0.3377644E+02 \end{array} \left. \vphantom{\begin{array}{ccc} 0.7967131E+00 \\ -0.2937448E+01 \\ 0.6891613E+01 \end{array}} \right\} \underline{C} = [c_{\mu\nu}]$$

$\mu \uparrow$
 $\nu \leftarrow$

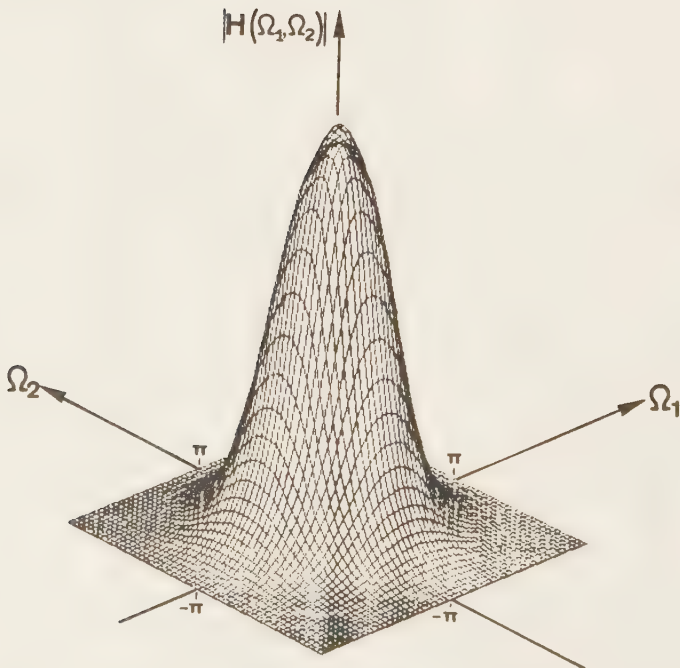
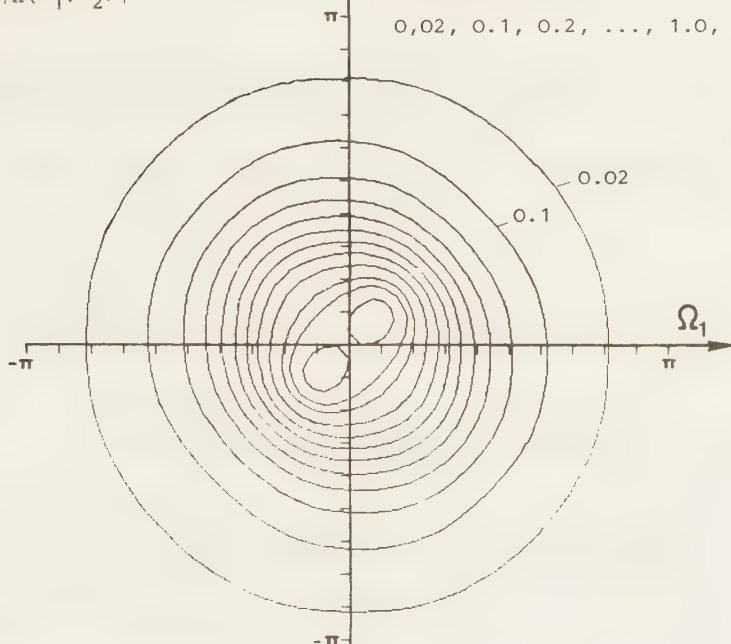
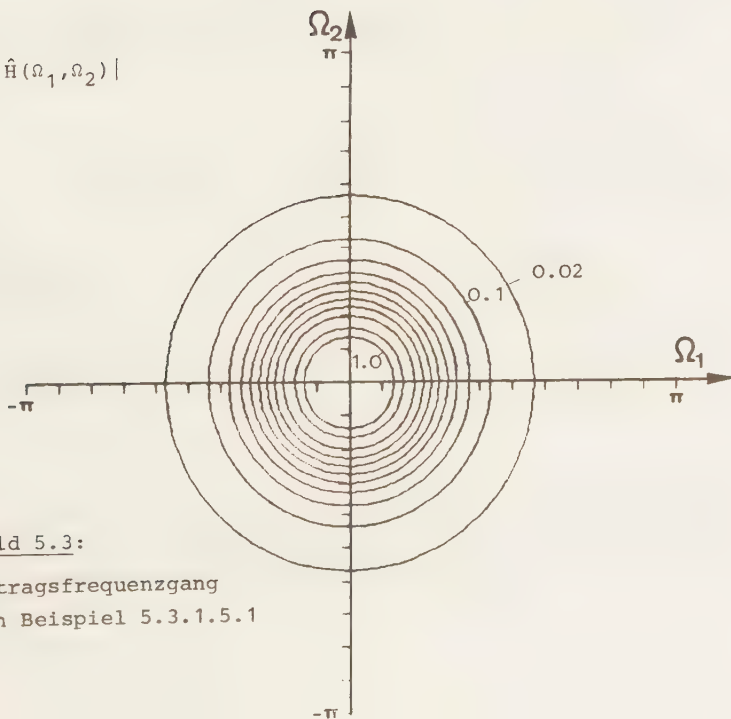


Bild 5.2: Betragsfrequenzgang und Koeffizientenmatrix des 2-D Filters von Beispiel 1

a) $|H(\Omega_1, \Omega_2)|$

Höhenlinienabstand:

0,02, 0.1, 0.2, ..., 1.0, 1.05

b) $|\hat{H}(\Omega_1, \Omega_2)|$ Bild 5.3:

Betragsfrequenzgang
von Beispiel 5.3.1.5.1

sche Verlauf bei diesem geringen Aufwand schon relativ gut, besonders im Vergleich zu bekannten Verfahren [28,50,83, 84].

Durch Kaskadierung zweier entgegengesetzt rekursierender Filter

$$\hat{H}(z_1, z_2) = H(z_1, z_2) \cdot H(z_1, z_2^{-1}) \quad (5.3-49)$$

kann die Zirkularsymmetrie wesentlich verbessert werden (Bild 5.3b), zudem erhält man dadurch bzgl. z_2 Nullphase (Kap.4.3.1).

5.3.1.5.2 Beispiel 2: Elliptisches Filter

Als zweites Beispiel wird ein Tiefpaß nach den Anforderungen vom ASSP-Workshop 1979 [85] entworfen.

Gefordert ist eine Durchlaßgrenze $\Omega_D = 0,36\pi$, eine Sperrgrenze $\Omega_S = 0,5\pi$, die Abweichungen im Durchlaßbereich sollen maximal $\pm 5\%$, im Sperrbereich $\pm 1\%$ betragen.

Als eindimensionales Prototypfilter wurde ein elliptisches Filter 4.Grades (Cauer-Tiefpaß) mit den Koeffizienten:

KOEFFIZIENTEN DES 1D FILTERS

BLOCKANZAHL = 2

KONSTANTE = 0.1099617E+00

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1236699E+00	C1 =	-0.5254451E+00
B0 =	0.1000000E+01	C0 =	0.7247763E+00

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1387520E+01	C1 =	-0.5272255E+00
B0 =	0.1000000E+01	C0 =	0.1855627E+00

entworfen.

Für die Zählerkoeffizienten der beiden Teilblöcke wird Quadrantensymmetrie gefordert (5.3-21). Damit ergeben sich insgesamt 6 freie Parameter, die noch optimiert werden müssen.

Die 2-D Optimierung führt auf ein 2-D Filter mit den Koeffizienten:

KOEFFIZIENTEN DES 2D FILTERS

KONSTANTE = 0.1099617E+00

0.5160889E+01	0.5678222E+01	0.5160889E+01	↑ v	$B_1 = [b_{\mu v}]$
0.5678222E+01	-0.9377727E+01	0.5678222E+01		
0.5160889E+01	0.5678222E+01	0.5160889E+01		

← μ

0.5613730E+01	0.2494605E+00	0.5733231E+01	↑ v	$C_1 = [c_{\mu v}]$
0.2494605E+00	-0.5617329E+01	-0.3039254E+01		
0.5733231E+01	-0.3039254E+01	0.1330602E+02		

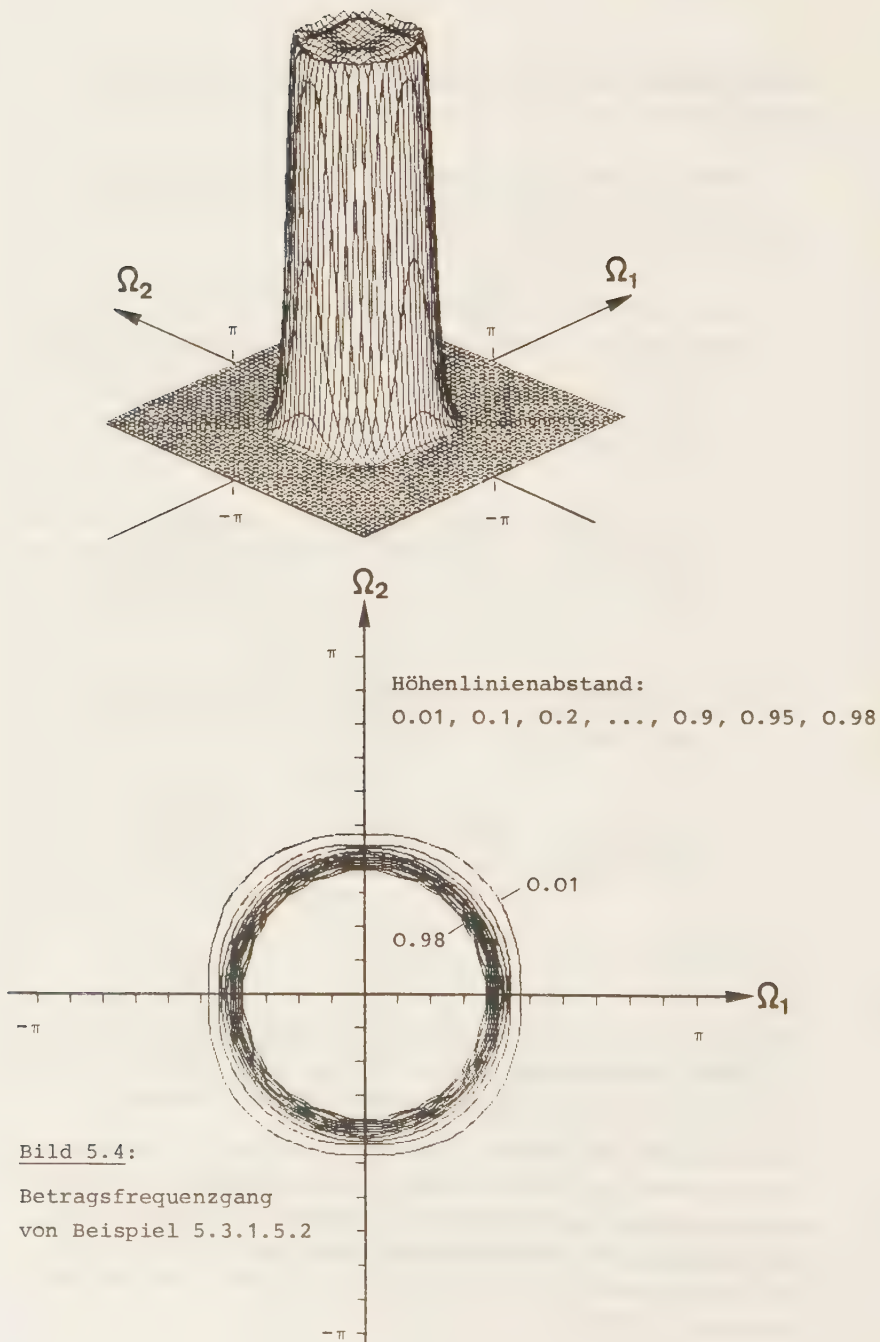
← μ

0.5883522E+01	0.4232956E+01	0.5883522E+01	} B_2
0.4232956E+01	0.1373441E+02	0.4232956E+01	
0.5883522E+01	0.4232956E+01	0.5883522E+01	

0.5247128E-01	-0.8886366E+00	0.3805169E+01	} C_2
-0.8886366E+00	0.9786236E+01	-0.1733321E+02	
0.3805169E+01	-0.1733321E+02	0.2952804E+02	

In Bild 5.4 sind der Betragsfrequenzgang dreidimensional und dessen Höhenlinien dargestellt, wobei zwei in + + und + - Richtung rekursierende Filter kaskadiert werden.

Mit diesem Filter lassen sich die Toleranzanforderungen übererfüllen. Die Abweichungen im Durchlaßbereich betragen weniger als 2%, im Sperrbereich weniger als 1%. Für die Realisierung sind 30 Multiplikationen und 64 Additionen pro Ausgangswert notwendig, sofern die Koeffizientensymmetrien ausgenutzt werden.



Zusammenfassend kann gesagt werden, daß diese neue Entwurfsmethode erlaubt, stabile rekursive Filter mit guter Zirkularsymmetrie zu entwerfen. Dabei wird der Entwurfs- und Optimierungsaufwand gegenüber direkten zweidimensionalen Entwurfsverfahren erheblich reduziert. Vor dem Entwurf eines eindimensionalen Prototypfilters wird der minimale Filtergrad berechnet. Damit wird die bei anderen Verfahren auftretende Unsicherheit bei der Festlegung des Filtergrades und der Zuordnung der Dimension von Zähler- und Nennerkoeffizientenmatrix vermieden. Außerdem ist die Stabilität des zweidimensionalen Filters stets garantiert.

Das vorgestellte Verfahren ist nicht nur auf zirkularsymmetrische Filter beschränkt. So kann durch Änderung der Bedingungen (5.3-10) und (5.3-11) erreicht werden, daß das zweidimensionale Filter auf den Hauptachsen $(\Omega_1, 0)$, $(\Omega_2, 0)$ den Frequenzgang zweier unterschiedlicher eindimensionaler Filter besitzt. Beispielsweise können (5.3-10) und (5.3-11) durch die Bedingungsgleichungen

$$G_2(w_1, 0) \stackrel{!}{=} G_1(w_1)$$

und $G_2(0, w_2) \stackrel{!}{=} \tilde{G}_1(w_2)$

ersetzt werden, wobei $G_1(w)$, $\tilde{G}_1(w)$ die Übertragungsfunktionen zweier unterschiedlicher 1-D Prototypfilter sind.

Im Gegensatz zur Spektraltransformation mit einer Allpaß-Übertragungsfunktion $|50|$ tritt hier eine Periodizität des Frequenzgangs bei $(\Omega_1, \Omega_2) = (\pi, \pi)$ nicht auf. Da Zähler- und Nenner getrennt transformiert werden können, hat man mehr Freiheitsgrade für die Optimierung eines bestimmten Wunschverlaufs als bei bekannten Spektraltransformationsverfahren $[50, 72, 83, 84]$. Dies führt auf Systeme mit niedrigerem Systemgrad, und als Folge davon reduziert sich der Realisierungsaufwand. Wie bei allen anderen Transformationsverfahren ist auch bei diesem Verfahren nicht von vornherein

garantiert, daß das Toleranzschema durch das zweidimensionale Filter sicher eingehalten wird, wenn das entsprechende 1-D-Toleranzschema vom 1-D Prototypfilter erfüllt wird.

Diesen Nachteil vermeidet das im nächsten Abschnitt beschriebene Entwurfsverfahren.

5.3.2 Entwurfsverfahren 2 | 86 |

Ziel dieses Verfahrens ist der Entwurf von zweidimensionalen rekursiven digitalen Filtern mit näherungsweise zirkularsymmetrischem Betragsfrequenzgang ausgehend von 1-D-Prototypfiltern, wobei garantiert wird, daß das vorgegebene Toleranzschema eingehalten wird. Eine 2-D-Optimierung ist im Gegensatz zu Entwurfsverfahren 1 nicht erforderlich. Das resultierende zweidimensionale Filter wird aus verschiedenen gedrehten eindimensionalen Filtern erzeugt und ist stabil, sofern die 1-D Filter stabil sind. Bereits in [74] wurde ein ähnliches Verfahren beschrieben. Auf die dort nötige Interpolation kann aber verzichtet werden. Weiterhin wird untersucht, welchen Restriktionen die Bandgrenzen unterliegen und welche Forderungen an die Toleranzen der 1-D Prototypfilter zu stellen sind, damit ein vorgegebenes 2-D Toleranzschema sicher eingehalten wird.

Das Verfahren wird am Beispiel eines 2-D digitalen IIR-Tiefpasses erläutert. Analog dazu läßt es sich auch auf andere Filtertypen (Hochpaß, Bandpaß, Bandsperre) anwenden.

5.3.2.1 Drehung eines eindimensionalen Filters

Durch zeilenweise Anwendung eines 1-D Prototyptiefpasses mit dem (idealen) Betragsfrequenzgang

$$|H_O(\Omega)| = \begin{cases} 1 & 0 \leq |\Omega| \leq \Omega_D \\ 0 & \Omega_S \leq |\Omega| < \pi \end{cases}, \quad (5.3-50)$$

$$\Omega = 2\pi \frac{f}{f_A}, \quad f_A = \text{Abtastfrequenz}$$

$$\begin{aligned} \text{und } \Omega_D &= \text{Durchlaßgrenze,} \\ \Omega_S &= \text{Sperrgrenze} \end{aligned}$$

auf ein zweidimensionales Eingangssignal (Bild 5.5b) ergibt sich der in Bild 5.5c dargestellte Betragsfrequenzgang $|H_I(\Omega_1, \Omega_2)|$ des dadurch realisierten 2-D Filters. Analytisch entspricht dies einer Spektraltransformation der Variablen Ω bzw. $z = e^{j\Omega}$ mit $\Omega = \Omega_1$, oder

$$z = z_1, \quad (5.3-51)$$

weshalb

$$H_O(z) = H_O(z_1) = H_I(z_1, z_2) \quad (5.3-52)$$

wird. Der Betragsfrequenzgang des 1-D-Filters ist mit 2π periodisch (Bild 5.5a), die gleiche Periodizität ergibt sich dann wegen (5.3-51) auch für das 2-D-Filter bzgl. Ω_1 (Bild 5.5c).

Die Anwendung der eindimensionalen Filter in Diagonalrichtung nach Bild 5.6a wird durch die Transformation

$$z = z_1 \cdot z_2 \quad (5.3-53)$$

erreicht. Damit erhält man einen um $\beta=45^\circ$ gedrehten Betragsfrequenzgang des zweidimensionalen Filters

$$H(z) = H(z_1 \cdot z_2) = H_{II}(z_1, z_2) \quad (5.3-54)$$

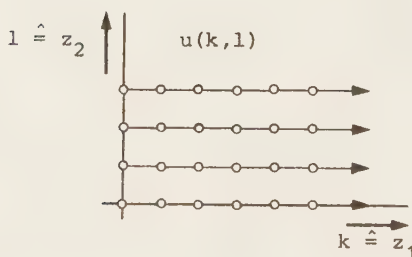
(Bild 5.6b). Allgemein entspricht die Anwendung eines 1-D Filters in bestimmter Richtung, beschrieben durch den Drehwinkel β , einer Spektraltransformation der Form

$$z = z_1^r \cdot z_2^q, \quad (5.3-55)$$

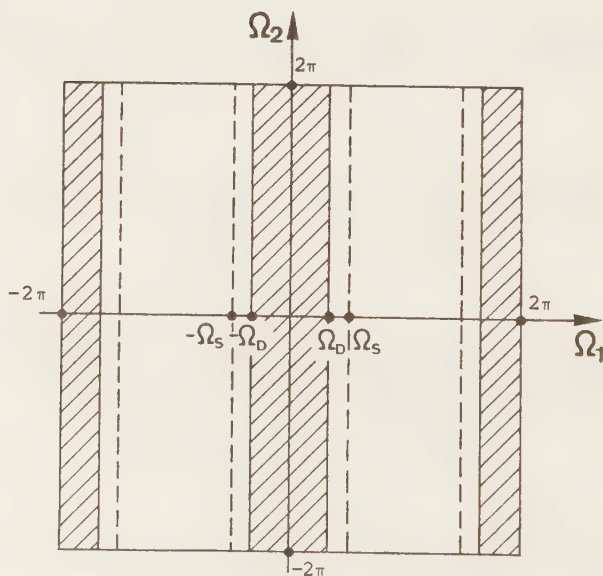
wobei r, q ganze oder rationale Zahlen sind.



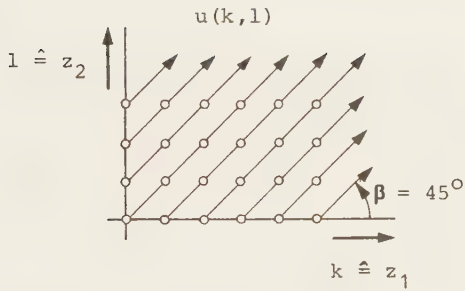
a) Betragsfrequenzgang des 1-D-Prototypfilters



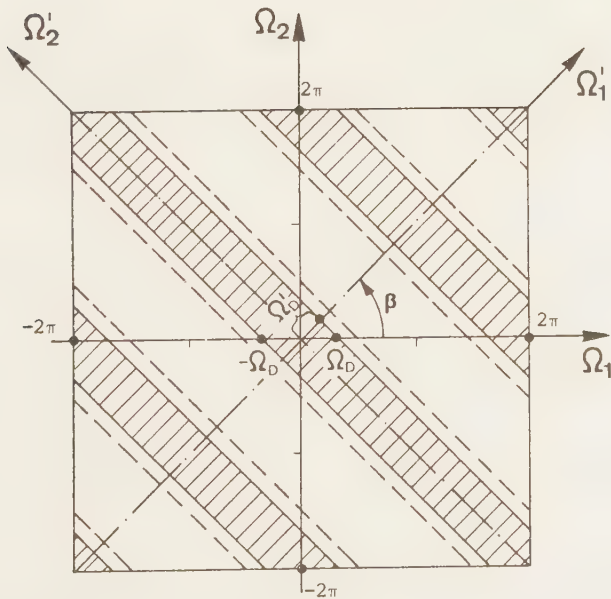
b) Zeilenweise Rekursion des Eingangsbildes



c) Betragsfrequenzgang $|H_I(\Omega_1, \Omega_2)|$



a) Rekursion in Diagonalrichtung



b) Betragsfrequenzgang $|H_{II}(\Omega_1, \Omega_2)|$

Bild 5.6: Frequenzgang des um 45° gedrehten Filters

Der Drehwinkel β berechnet sich zu

$$\beta = \arctan \frac{q}{r} . \quad (5.3-56)$$

Für $r > 0$, $q > 0$ ergibt sich ein kausales 2-D-Filter. Hier ist anzumerken, daß es gleichwertig ist, entweder das Eingangssignal mit einem um β gedrehten Filter zu verarbeiten, oder das um $-\beta$ gedrehte Eingangssignal mit einem ungedrehten Filter zu filtern.

Falls q und r in (5.3-55) ganzzahlig sind, liegen die erforderlichen Punkte des Eingangssignals wieder auf dem ursprünglichen Punktraster, andernfalls ist eine Interpolation des Eingangsbildes zur Berechnung der Zwischenpunkte notwendig [74].

Da nach der Drehung um β

$$\Omega_1' = \Omega_1 \cdot \cos \beta$$

gilt (Bild 5.6b), werden die Bandgrenzen kontrahiert. So ergibt sich die Durchlaßgrenze des gedrehten Filters zu

$$\Omega_D' = \Omega_D \cdot \cos \beta , \quad (5.3-57)$$

entsprechendes gilt für die Sperrgrenze. Dies ist beim Entwurf des 1-D Prototypfilters mit zu berücksichtigen.

Ein weiterer Effekt ist noch wichtig. Da der Frequenzgang des 1-D-Filters mit 2π periodisch ist, ergibt sich nach der Drehung ein längs der Ω_1' -Achse (Bild 5.6b) mit $2\pi \cdot \cos \beta$ periodischer Frequenzgang. Dadurch fällt ein Teil des Durchlaßbereiches der zweiten Periode in das gewünschte Sperrintervall. Dieser Effekt kann durch Übergang auf ein rationales r in (5.3-55) vermieden bzw. verringert werden. Das entspricht einer dichteren Abtastung der Eingangsfolge, bzw. macht eine Interpolation des diskreten Eingangssignals erforderlich.

5.3.2.2 Kaskadierung von um $v \cdot \frac{\pi}{4}$ gedrehten 1-D-Filtern

In diesem Abschnitt wird gezeigt, wie sich Kaskadierung von vier um $v \cdot \frac{\pi}{4}$ $v \in \{-1, 0, 1, 2\}$ gedrehten 1-D-Filtern 2-D-Filter mit näherungsweise zirkulärsymmetrischem Betragsfrequenzgang entwerfen lassen.

Gewünscht sei ein digitaler 2D-IIR-TP mit zirkulärsymmetrischem Betragsfrequenzgang, d.h.

$$| |H(\Omega_1, \Omega_2)| - 1 | \leq \delta_1 \quad \text{für} \quad \sqrt{\Omega_1^2 + \Omega_2^2} \leq \Omega_D < \Omega_S$$

und

$$|H(\Omega_1, \Omega_1)| \leq \delta_2 \quad \text{für} \quad \Omega_S \leq \sqrt{\Omega_1^2 + \Omega_2^2} \leq \frac{2\pi}{2 + \sqrt{2}} \quad (5.3-58)$$

Entworfen werden zunächst zwei 1-D-Prototypfilter mit den Übertragungsfunktionen $H_1(z)$ und $H_2(z)$, wobei folgende Spezifikationen zu erfüllen sind:

- Teilfilter 1, $H_1(z)$:

$$\begin{aligned} |1 - |H_1(\Omega)| | &\leq \delta_1^{(1)} & |\Omega| &\leq \Omega_D \\ |H(\Omega)| &\leq \delta_2^{(2)} & \Omega_S &\leq |\Omega| \leq \pi \end{aligned} \quad (5.3-59a)$$

- Teilfilter 2, $H_2(z)$:

$$\begin{aligned} |1 - |H_2(\Omega)| | &\leq \delta_1^{(2)} & |\Omega| &\leq \sqrt{2} \Omega_D \\ |H_2(\Omega)| &\leq \delta_2^{(2)} & \sqrt{2} \Omega_S &\leq |\Omega| \leq \pi. \end{aligned} \quad (5.3-59b)$$

Wendet man das erste 1-D-Teilfilter nacheinander auf die Zeilen und Spalten an, so erhält man eine separable 2-D-Übertragungsfunktion eines Teilsystems a

$$H_a(z_1, z_2) = H_1(z_1) \cdot H_1(z_2) \quad , \quad (5.3-60)$$

dessen Durchlaßgrenze näherungsweise ein Quadrat mit der Kantenlänge $2 \Omega_D$ ist (Bild 5.7a). Das zweite 1-D-Teilfilter wenden wir nacheinander auf die Abtastwerte in beiden Diagonalrichtungen an. Das ergibt

$$H_b(z_1, z_2) = H_2(z_1 \cdot z_2) \cdot H_2(z_1 \cdot z_2^{-1}). \quad (5.3-61)$$

Die Durchlaßgrenze dieses Filters beschreibt näherungsweise ein um 45° gedrehtes Quadrat mit der Kantenlänge $2 \cdot \Omega_D$. Wegen der gewählten Transformation ist der Frequenzgang mit (π, π) periodisch, was hier jedoch nicht weiter störend ist (Bild 5.7b).

Durch Kaskadierung der beiden zweidimensionalen Teilfilter a und b ergibt sich schließlich die Übertragungsfunktion $H(z_1, z_2)$ des gewünschten 2-D IIR-Filters, dessen Durchlaßgrenze durch ein gleichseitiges Achteck beschrieben wird (Bild 5.7c). Je steiler der Übergangsbereich des Filters ist, desto genauer wird dieser Verlauf eingehalten. Dies stellt für praktische Anwendungen eine hinreichend gute Approximation eines zirkulärsymmetrischen Betragsfrequenzgangs dar. Alle 1-D Teilfilter verwenden für die Filterung ausschließlich Abtastwerte auf dem ursprünglichen Punktraster.

Durch Kaskadierung von paarweise entgegengesetzt rekursierenden identischen 1-D Filtern lassen sich auf einfache Weise Nullphasenfilter realisieren.

5.3.2.3 Einschränkung der Sperrgrenze

Für das 2-D Teilsystem a ergibt sich keine Kontraktion der Bandgrenzen, d.h. es ist

$$\Omega_D^{(1)} = \Omega_D, \quad \Omega_S^{(1)} = \Omega_S. \quad (5.3-62)$$

Das Teilsystem b setzt sich aus um $\pm 45^\circ$ gedrehten 1-D Filtern zusammen, nach (5.3-56) werden demnach die Bandgrenzen

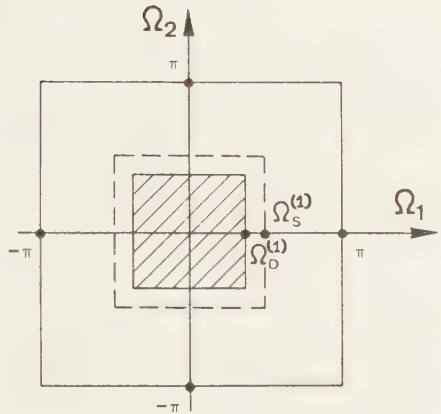
a)

Drehung um $0^\circ, 90^\circ$ ($v=0,2$):
keine Kontraktion der Band-
grenzen

1-D Teilfilter 1: $H_1(z)$

$$\Omega_D^{(1)} = \Omega_D, \quad \Omega_S^{(1)} = \Omega_S$$

$$H_a(z_1, z_2) = H_1(z_1) \cdot H_1(z_2)$$



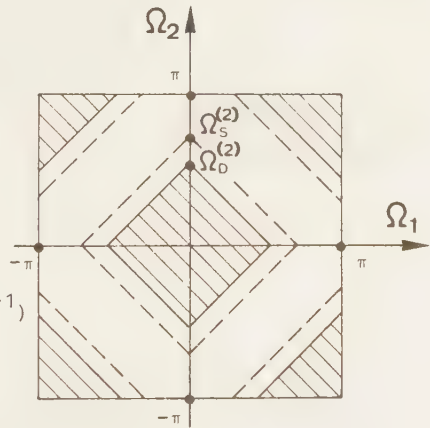
b)

Drehung um $\pm 45^\circ$ ($v=\pm 1$)
Kontraktion der Bandgrenzen
um $1/\sqrt{2}$

1-D Teilfilter 2: $H_2(z)$

$$\Omega_D^{(2)} = \sqrt{2} \Omega_D, \quad \Omega_S^{(2)} = \sqrt{2} \Omega_S$$

$$H_b(z_1, z_2) = H_2(z_1 \cdot z_2) \cdot H_2(z_1 \cdot z_2^{-1})$$



c)

2-D Filter

$$H(z_1, z_2) = H_a(z_1, z_2) \cdot H_b(z_1, z_2)$$

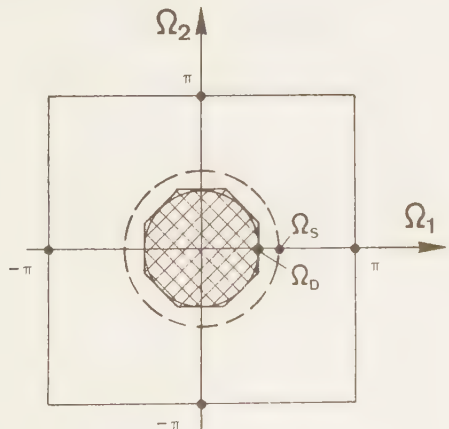


Bild 5.7: Kaskadierung von 4 um $v \cdot \frac{\pi}{4}$ gedrehten 1-D Filtern

um $\sqrt{2}$ kontrahiert, weshalb das 1-D Filter mit

$$\Omega_D^{(2)} = \frac{\Omega_D}{\cos\beta} = \Omega_D \cdot \sqrt{2} \quad (5.3-63a)$$

$$\Omega_S^{(2)} = \frac{\Omega_S}{\cos\beta} = \Omega_S \cdot \sqrt{2} \quad (5.3-63b)$$

entworfen werden muß.

Aus der Forderung, daß die Übergangsbereiche der Filter a und b nicht überlappen sollen, ergibt sich die Bedingungs-
gleichung

$$\Omega_S^{(1)} \cdot \sqrt{2} + \frac{\Omega_S^{(2)}}{\sqrt{2}} \leq \pi \cdot \sqrt{2} \quad (5.3-64)$$

$\Omega_S^{(1)}$, $\Omega_S^{(2)}$ nach (5.3-58) (5.3-59) eingesetzt, führt auf
die Einschränkung

$$\Omega_S \leq \frac{2\pi}{2+\sqrt{2}} \approx 0,587\pi \quad (5.3-65)$$

Im nächsten Abschnitt soll untersucht werden, welche Anforderungen an die Toleranzen der 1-D-Prototypfilter zu stellen sind, damit ein vorgegebenes Toleranzschema für das zweidimensionale Filter nicht verletzt wird.

5.3.2.4 Anforderungen an die Toleranzen der 1-D-Filter

Durch die Kaskadierung der um $\nu \cdot \frac{\pi}{4}$ gedrehten Teilfilter ergeben sich 4 Bereiche, in denen die Einhaltung der geforderten Toleranzen besonders zu überprüfen ist. Diese sind in Bild 5.8 dargestellt.

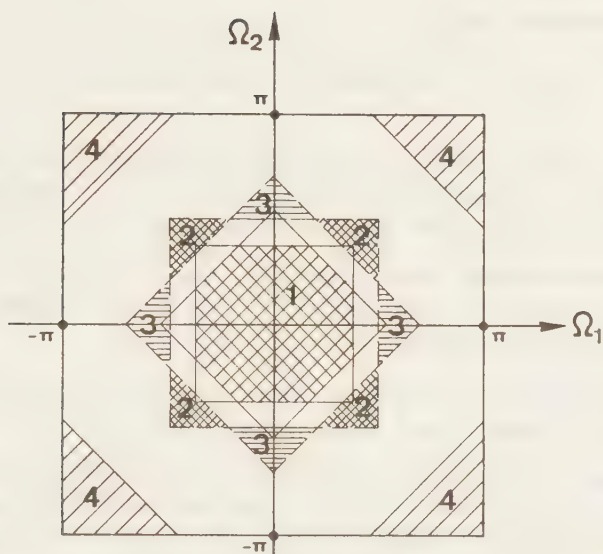


Bild 5.8: Kritische Bereiche für die Einhaltung der Toleranzen

Damit das vorgegebene Toleranzschema für das resultierende 2-D-Filter (5.3-58) sicher eingehalten wird, müssen von den Toleranzen der 1-D-Filter $\delta_1^{(v)}, \delta_2^{(v)}, v \in \{1, 2\}$ (5.3-59a,b) folgende 4 Ungleichungen erfüllt werden:

für Bereich 1:

$$1 - \delta_1 \leq (1 - \delta_1^{(1)})^2 (1 - \delta_1^{(2)})^2 \quad (5.3-66)$$

und
$$(1 + \delta_1^{(1)})^2 (1 + \delta_1^{(2)})^2 \leq 1 + \delta_1, \quad (5.3-67)$$

für Bereich 2:

$$(1 + \delta_1^{(1)})^2 (1 + \delta_1^{(2)}) \delta_2^{(2)} \leq \delta_2, \quad (5.3-68)$$

für Bereich 3 und 4:

$$(1+\delta_1^{(2)})^2 (1+\delta_1^{(1)}) \delta_2^{(1)} \leq \delta_2^{(1)} \quad (5.3-69)$$

Dabei wurde der ungünstigste Fall angenommen, daß im Bereich 2,3,4 jeweils 3 Durchlaßbereiche und 1 Sperrbereich der gedrehten 1-D-Filter zusammenfallen.

Bei vorgegebenen Toleranzen δ_1, δ_2 des 2-D-Filters lassen sich daraus die Toleranzen $\delta_1^{(v)}, \delta_2^{(v)}$ $v \in \{1,2\}$ der 1-D-Prototypfilter berechnen.

Wie sich leicht nachweisen läßt, wird Ungl.(5.3-66) stets erfüllt, sofern Ungl.(5.3-67) gilt.

Da der Übergangsbereich für das 1-D-Teilfilter 2 wegen (5.3-6a,b) um $\sqrt{2}$ breiter ist als für das Teilfilter 1, können die Toleranzvorschriften von Teilfilter 2 leichter erfüllt werden. Deshalb wird

$$\begin{aligned} \delta_1^{(1)} &\geq \delta_1^{(2)} \\ \delta_2^{(1)} &\geq \delta_2^{(2)} \end{aligned}$$

angenommen. Für $\delta_1^{(1)} = \delta_1^{(2)}$ ergibt sich aus (5.3-67)

$$\delta_1^{(2)} \leq \sqrt[4]{1+\delta_1^{(2)}} - 1 \quad (5.3-70)$$

Dies in (5.3-68) eingesetzt, führt zu

$$\delta_2^{(2)} \leq \frac{\delta_2^{(2)}}{(1+\delta_1^{(2)})^3} \quad (5.3-71)$$

Mit (5.3-70),(5.3-71) hat man einen Anhaltspunkt, wie die Toleranzen für den Entwurf des Teilfilters 2 vorzuschreiben sind. Wird zunächst das Teilfilter 2 entworfen, kennt man nach dem Entwurf die tatsächlichen maximalen Abweichungen

im Durchlaß- $(\delta_{1a}^{(2)})$ und Sperrbereich $(\delta_{1a}^{(2)})$. Damit lassen sich anschließend die Toleranzvorschriften für den Entwurf von Teilfilter 1 berechnen.

Aus (5.3-67), (5.3-68) erhält man

$$\delta_1^{(1)} \leq \sqrt{\text{Min} \left\{ \frac{\delta_2}{(1+\delta_{1a}^{(2)}) \delta_{2a}^{(2)}}, \frac{1+\delta_1^{(2)}}{(1+\delta_{1a}^{(2)})^2} \right\}} - 1 \quad (5.3-72)$$

und aus (5.3-69)

$$\delta_2^{(1)} \leq \frac{\delta_2}{(1+\delta_{1a}^{(2)})^2 (1+\delta_1^{(1)})}, \quad (5.3-73)$$

wobei sich $\delta_1^{(1)}$ nach (5.3-72) berechnet. Für den Fall, daß ein Nullphasenfilter realisiert werden soll, sind die Toleranzen δ_1, δ_2 entsprechend umzurechnen. Für die Nullphasenkaskade (Kap.4.3.1) ergibt sich

$$\delta_1 \leq \sqrt{1+\delta_{10}} - 1 \quad (5.3-74)$$

$$\text{und} \quad \delta_2 \leq \sqrt{\delta_{20}}$$

mit δ_{10}, δ_{20} als Toleranzen des Nullphasenfilters.

Durch die in diesem Abschnitt beschriebene Vorgehensweise können die Toleranzvorschriften der 1-D-Prototypfilter aus den Toleranzvorschriften des gewünschten 2-D-Filters berechnet werden. Damit ist garantiert, daß das Toleranzschema des 2-D-Filters eingehalten wird. Da die Bandgrenzen der 1-D-Filter nach (5.3-62) und (5.3-63) ebenfalls festliegen, kann auch der dazu notwendige minimale Filtergrad $|1|$ bestimmt werden. Die sonst übliche Unsicherheit der Wahl des Filtergrades beim Entwurf von 2-D-Filtern wird damit auch hier vermieden.

5.3.2.5 Verallgemeinerung

Das beschriebene Verfahren ist nicht auf die Kaskadierung von $v \cdot \frac{\pi}{4}$ gedrehten 1-D-Filtern beschränkt, obwohl dieser Fall in Bezug auf den notwendigen geringen Realisierungsaufwand und der relativ guten Approximation eines zirkularsymmetrischen Betragsfrequenzgangs optimal ist.

Es sei jetzt verallgemeinernd angenommen, daß das zweidimensionale TP-Filter durch Kaskadierung von Teilsystem "a" nach (5.3-60) mit verschiedenen Teilsystemen "b_μ" μ=1,...,m erzeugt wird, d.h.

$$H(z_1, z_2) = H_a(z_1, z_2) \cdot \prod_{\mu=1}^m H_{b_\mu}(z_1, z_2) , \quad (5.3-75a)$$

$$\text{wobei } H_{b_\mu}(z_1, z_2) = H_\mu(z_1^{r_\mu} \cdot z_2^{q_\mu}) \cdot H_\mu(z_1^{r_\mu} z_2^{-q_\mu}) \quad (5.3-75b)$$

ist. Dabei sind r_μ, q_μ im allgemeinen rationale Zahlen.

Untersucht werden soll nun, welche Bedingungen die Durchlaßgrenze Ω_D und die Sperrgrenze Ω_S des zweidimensionalen Gesamtsystems erfüllen müssen, damit es in dieser Weise realisierbar ist. Dazu wird die Kaskade

$$\begin{aligned} & H_a(z_1, z_2) \cdot H_b(z_1, z_2) \\ &= H_1(z_1) \cdot H_1(z_2) \cdot H(z_1^{r_1} \cdot z_2^{q_1}) \cdot H(z_1^{r_1} \cdot z_2^{-q_1}) \end{aligned} \quad (5.3-76)$$

betrachtet, wobei $H_b(z_1, z_2)$ gerade das Teilsystem "b_μ" mit dem größten $|q_\mu|$ und falls mehrere vorhanden sind, das mit dem größten Winkel

$$\beta_0 = \arctan \left| \frac{r_\mu}{q_\mu} \right| \text{ ist.} \quad (5.3-77)$$

Zur Ableitung der Bedingung für die Sperrgrenze betrachten wir Bild 5.8a

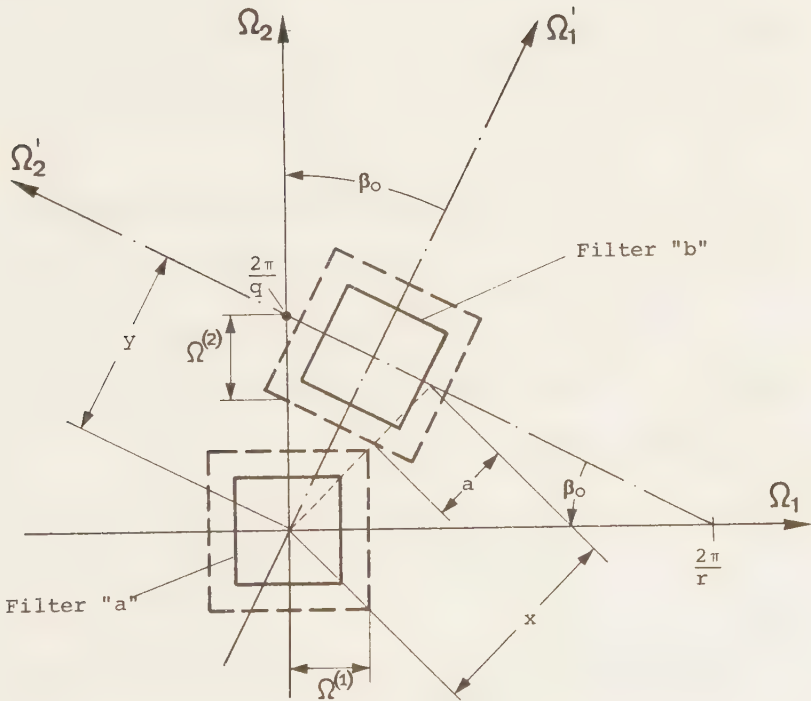


Bild 5.8a: Zur Ableitung der Bedingung für die Sperrgrenze Ω_S

Ohne Beschränkung der Allgemeinheit wurde dabei $0 < r < q$ vorausgesetzt. Wegen der Übersichtlichkeit ist für das Teilsystem "b" nur die 1. periodische Wiederholung des Betragsfrequenzgangs eingezeichnet.

Gefordert wird, daß die Übergangsbereiche von Filter "a" und "b" nicht überlappen. Dies ist erfüllt, wenn

$$\Omega_2^{(1)} \cdot \sqrt{2} + a \leq x \quad (5.3-78)$$

gilt.

$$\text{Mit } y = \frac{2\pi}{q} \cos \beta_0 ,$$

$$x = y / \cos (\pi/4 - \beta_0)$$

ergibt sich unter Anwendung des Strahlensatzes die Bedingung

$$\Omega_S^{(1)} (1 + \tan \beta_0) + \Omega_S^{(2)} \leq \frac{2\pi}{|q|} . \quad (5.3-79)$$

Da aber nach (5.3-62) $\Omega_S^{(1)} = \Omega_S$ und nach (5.3-63)

$$\Omega_S^{(2)} = \frac{\Omega_S}{\cos \beta_0} \text{ ist, ergibt sich}$$

$$\Omega_S \cdot (1 + \tan \beta_0) + \frac{\Omega_S}{\cos \beta_0} \leq \frac{2\pi}{|q|} .$$

Daraus folgt

$$\Omega_S \leq \frac{2\pi}{|q| (1 + \tan \beta_0 + \frac{1}{\cos \beta_0})} , \quad (5.3-80)$$

außerdem muß

$$\Omega_D < \Omega_S \quad (5.3-81)$$

gelten. Für 2-D-Filter, die bzgl. Ω_D und Ω_S die Ungl. (5.3-80), (5.3-81) erfüllen, ist dieses Entwurfsverfahren anwendbar. Für $q=r=1$ ergibt sich aus (5.3-80) die Bedingung (5.3-65) als Spezialfall.

5.3.2.6 Beispiele

Es wird der Entwurf von zwei Tiefpässen mit $\Omega_D = 0.36\pi$, $\Omega_S = 0.5\pi$, $\delta_1 \leq 5\%$, $\delta_2 \leq 1\%$ (entsprechend den Anforderungen von [85]) vorgenommen.

Die 1-D-Prototypfilter werden als Cauerfilter mit dem Programm DORED1 [82] an einer PDP 11/45 entworfen (Entwurfszeit: einige Sekunden).

5.3.2.6.1 Zirkulärsymmetrischer Tiefpaß

Dieses Beispiel entspricht dem nach Abschnitt 5.3.1.5.2

Die 1-D-Prototypfilter mit den Koeffizienten

KONSTANTE = 0.6330000E-04

Teilfilter 1: Grad 5

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.7788790E-01	C1 =	-0.5735437E+00
B0 =	0.1000000E+01	C0 =	0.7871526E+00

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.8465357E+00	C1 =	-0.5944148E+00
B0 =	0.1000000E+01	C0 =	0.3557250E+00

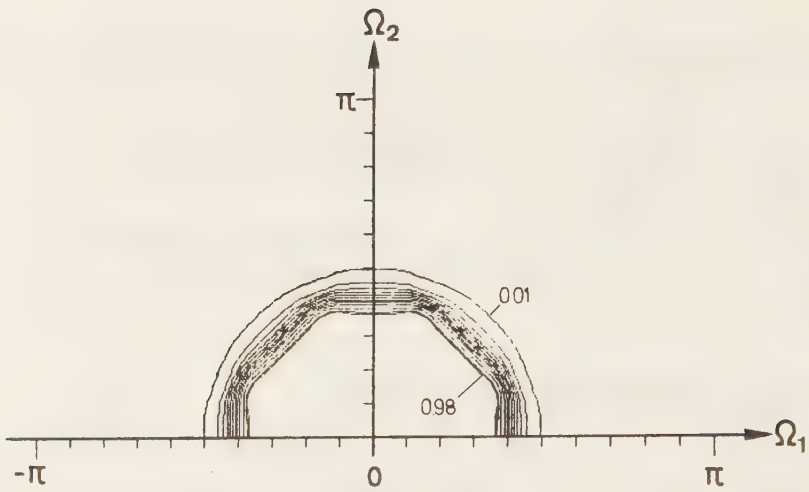
B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1000000E+01	C1 =	-0.3162201E+00
B0 =	0.0000000E+00	C0 =	0.0000000E+00

Teilfilter 2: Grad 4

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1294793E+01	C1 =	0.2535341E+00
B0 =	0.1000000E+01	C0 =	0.4883796E+00

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1843333E+01	C1 =	-0.1532482E+00
B0 =	0.1000000E+01	C0 =	0.1303674E+00

ergeben den in Bild 5.9 dargestellten Betragsfrequenzgang des zweidimensionalen Filters.



$$|H(\Omega_1, \Omega_2)| = 0.01, 0.1, 0.2, \dots, 0.9, 0.97, 0.98$$

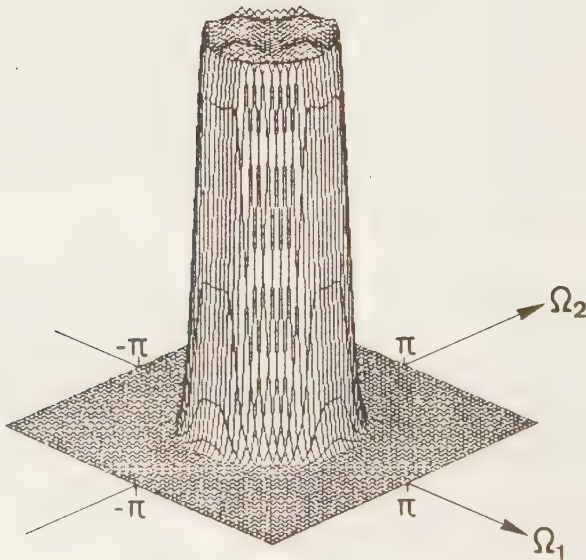


Bild 5.9: Betragsfrequenzgang $|H(\Omega_1, \Omega_2)|$ von 5.3.2.6.1

Für die Realisierung sind 27 Multiplikationen und 36 Additionen pro Ausgangswert erforderlich, also deutlich weniger als für Beispiel 5.3.1.5.2. Das liegt daran, daß nur 1-D-Filter für die Realisierung verwendet werden.

5.3.2.6.2 Zirkularsymmetrischer Tiefpaß mit Nullphase

Die 1-D-Filter ergeben sich zu

KONSTANTE = 0.5286460E-02

Teilfilter 1: Grad 4

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1236699E+00	C1 =	-0.3005367E+00
B0 =	0.1000000E+01	C0 =	0.7090070E+00
B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1387520E+01	C1 =	-0.9718731E-01
B0 =	0.1000000E+01	C0 =	0.1271885E+00

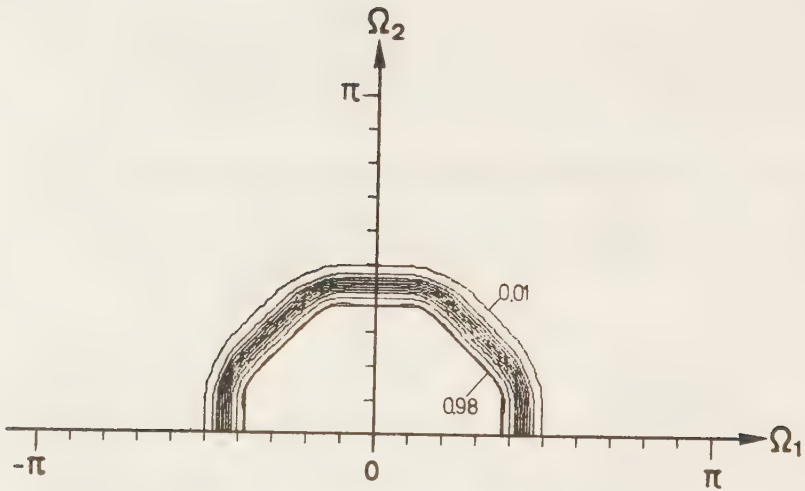
Teilfilter 2: Grad 3

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1358799E+01	C1 =	0.5220036E+00
B0 =	0.1000000E+01	C0 =	0.5574446E+00
B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1000000E+01	C1 =	0.1618818E+00
B0 =	0.0000000E+00	C0 =	0.0000000E+00

Den Betragsfrequenzgang des zweidimensionalen Filters zeigt Bild 5.10.

Auch hier werden die Toleranzen mit $\delta_1 < 2\%$, $\delta_2 < 1\%$ unterschritten.

Als Realisierungsaufwand sind 41 Multiplikationen und 56 Additionen pro Ausgangswert notwendig.



$$|H(\Omega_1, \Omega_2)| = 0.01, 0.1, 0.2, \dots, 0.9, 0.97, 0.98$$

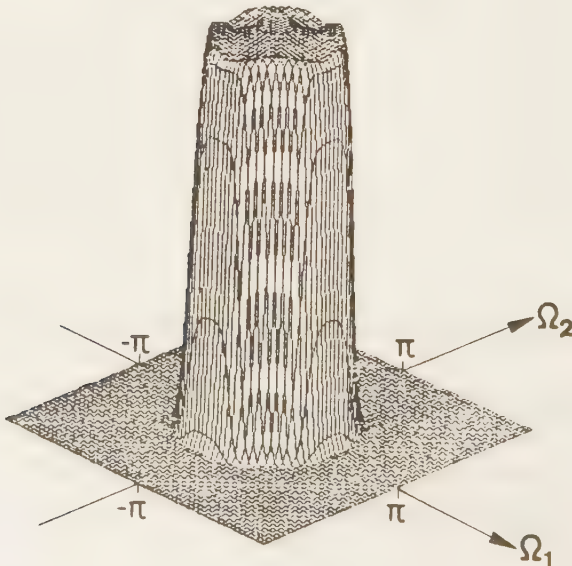


Bild 5.10: Betragsfrequenzgang des nullphasigen Filters nach 5.3.2.6.2

Im nächsten Abschnitt wird die Idee, 2-D-Filter durch Kadierung gedrehter 1-D-Filter zu entwerfen, auf den Entwurf von Fan-Filtern (z.B. [83-93]) angewendet.

5.3.3 Einfacher Entwurf von Fan-Filtern

Unter Fan-Filtern verstehen wir hier Filter mit fächerförmiger Übertragungsfunktion (Bild 5.11). Diese erlauben es zum Beispiel, Wellen, die entlang einer Geraden auf der Oberfläche eines Körpers gemessen werden, in Bereiche unterschiedlicher Fortpflanzungsgeschwindigkeit entlang dieser Geraden aufzugliedern [89]. Entsprechende Anwendungen sind in Zusammenhang mit geophysikalischen Untersuchungen bekannt [88-93].

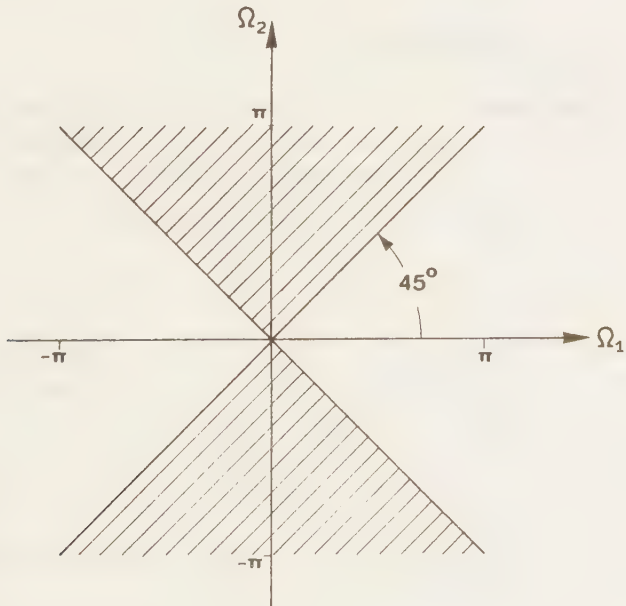


Bild 5.11: Frequenzgang eines idealen Fan-Filters
(Der Durchlaßbereich ist schraffiert)

Ein ideales Fan-Filter besitzt den Frequenzgang

$$H(\Omega_1, \Omega_2) = \begin{cases} 1 & , |\Omega_1| < |\Omega_2| \\ 0 & \text{sonst} \end{cases} \quad (5.3-82)$$

und die Impulsantwort [33]

$$h_o(k, l) = \begin{cases} \frac{1}{2} & , k=l=0 \\ \frac{2}{\pi^2(k^2-l^2)} & , k^2+l^2 \text{ ungerade} \\ 0 & \text{sonst} \end{cases} \quad (5.3-83)$$

5.3.3.1 Entwurfsprinzip

Bei dem Entwurf wird wieder von einem eindimensionalen Prototypfilter ausgegangen. Entworfen wird zunächst ein 1-D-Tiefpaß mit $\Omega_D = \frac{\pi}{2}$ und der Übertragungsfunktion $H_1(z)$.

Durch Anwendung der Spektraltransformation

$$H_1(z) \rightarrow H_1(-j z_1^{\frac{1}{2}} \cdot z_2^{\frac{1}{2}})$$

erhält man ein zweidimensionales Filter, dessen Frequenzgang den in Bild 5.12a dargestellten Durchlaßbereich besitzt. Für das transformierte Filter

$$H_1(z) \rightarrow H_1(j z_1^{\frac{1}{2}} \cdot z_2^{-\frac{1}{2}})$$

ergibt sich der in Bild 5.12b dargestellte Durchlaßbereich. Durch Kaskadierung dieser beiden Filter

$$H_o(z_1, z_2) = H_1(-j z_1^{\frac{1}{2}} \cdot z_2^{\frac{1}{2}}) \cdot H_1(j z_1^{\frac{1}{2}} \cdot z_2^{-\frac{1}{2}}) \quad (5.3-84a)$$

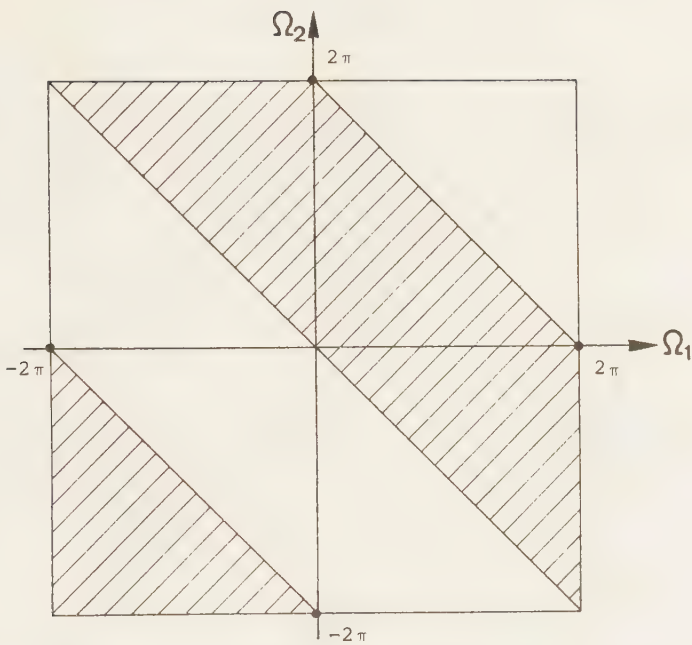


Bild 5.12a: Durchlaßbereich (schraffiert) von $H_1(-j z_1^{1/2} z_2^{1/2})$

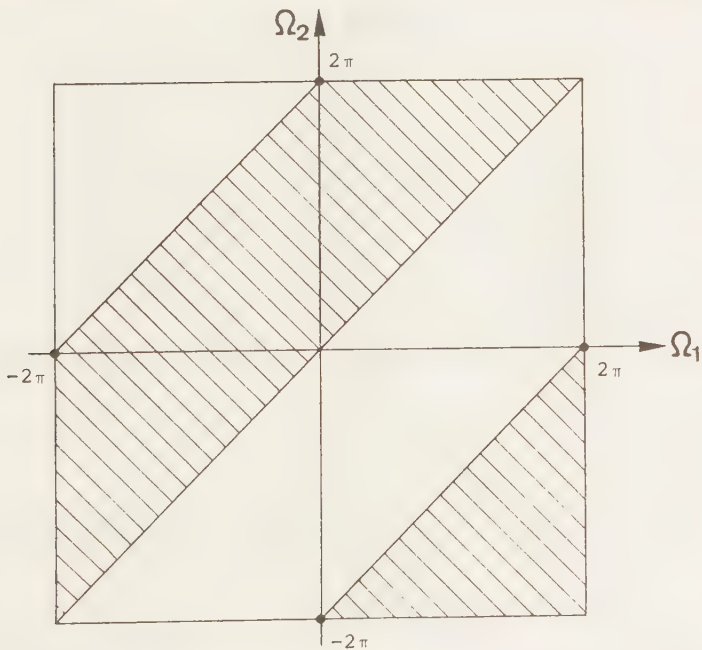


Bild 5.12b: Durchlaßbereich (schraffiert) von $H_1(j z_1^{1/2} z_2^{-1/2})$

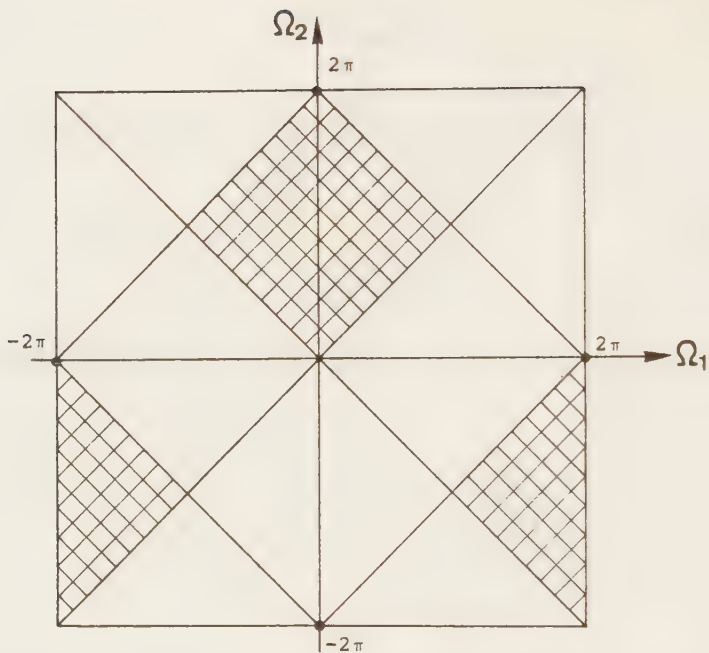


Bild 5.13a: Durchlaßbereich (schraffiert) von $H_0(z_1, z_2)$

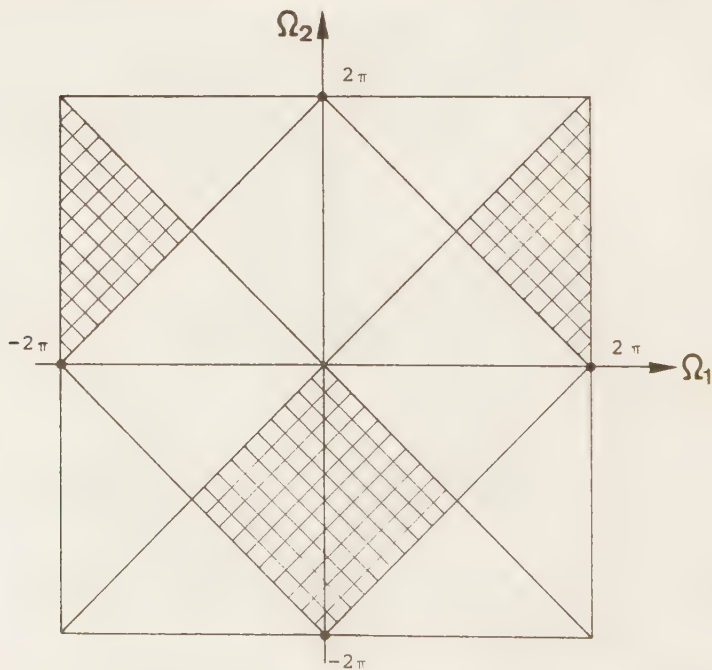


Bild 5.13b: Durchlaßbereich (schraffiert) von $\hat{H}_1(z_1, z_2)$

entsteht ein Teilfilter $H_O(z_1, z_2)$ mit komplexen Koeffizienten und dem in Bild 5.13a dargestellten Durchlaßbereich des Frequenzgangs. Analog dazu ergibt sich für

$$\hat{H}_O(z_1, z_2) = H_1\left(j z_1^{\frac{1}{2}} \cdot z_2^{\frac{1}{2}}\right) \cdot H_1\left(-j z_1^{\frac{1}{2}} \cdot z_2^{-\frac{1}{2}}\right) \quad (5.3-84b)$$

ein Durchlaßbereich nach Bild nach 5.13b.

Durch Kombination (Addition, Subtraktion) dieser beiden Teilfilter $H_O(z_1, z_2)$ und $\hat{H}_O(z_1, z_2)$ erhält man ein zweidimensionales Filter, dessen Betragsfrequenzgang dem eines Fan-Filters entspricht (Bild 5.14).

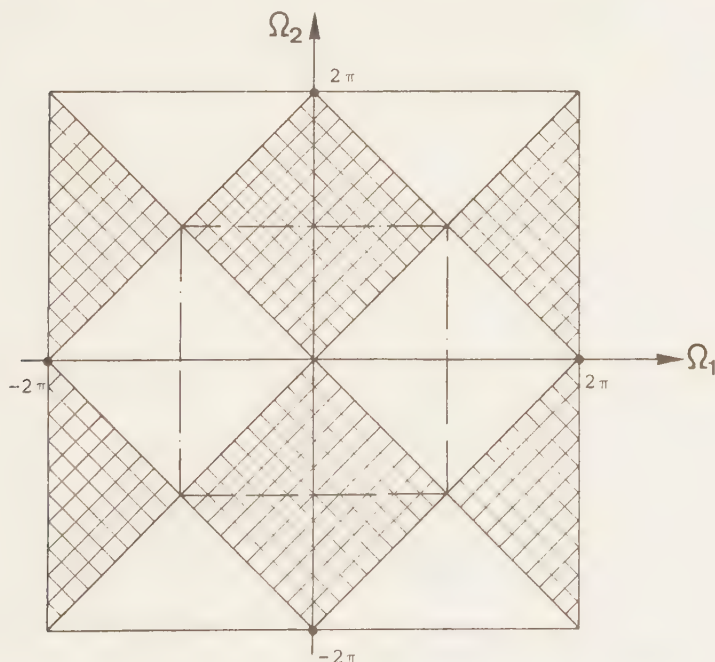


Bild 5.14: Betragsfrequenzgang von $H_O(z_1, z_2) + \hat{H}_O(z_1, z_2)$

Da nach (5.3-3) sich jedes 1-D-Filter als Kaskade von Blöcken 1. und 2. Grades darstellen läßt als

$$H(z) = \prod_{\lambda=1}^1 H_{\lambda}(z),$$

ergeben sich die zweidimensionalen Teilfilter $H_O(z_1, z_2)$, $\hat{H}_O(z_1, z_2)$ nach (5.3-84a,b) zu

$$\begin{aligned} H_O(z_1, z_2) &= \prod_{\lambda=1}^1 H_{\lambda}(-j z_1^{\frac{1}{2}} \cdot z_2^{\frac{1}{2}}) \cdot H_{\lambda}(j z_1^{\frac{1}{2}} \cdot z_2^{-\frac{1}{2}}) \\ &= \frac{\prod_{\lambda=1}^1 B_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}})}{\prod_{\lambda=1}^1 C_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}})} \end{aligned} \quad (5.3-85a)$$

und

$$\begin{aligned} \hat{H}_O(z_1, z_2) &= \prod_{\lambda=1}^1 H_{\lambda}(j z_1^{\frac{1}{2}} \cdot z_1^{\frac{1}{2}}) H_{\lambda}(-j z_1^{\frac{1}{2}} \cdot z_2^{-\frac{1}{2}}) \\ &= \frac{\prod_{\lambda=1}^1 \hat{B}_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}})}{\prod_{\lambda=1}^1 \hat{C}_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}})}, \end{aligned} \quad (5.3-85b)$$

wobei sich die Koeffizienten der Polynome $B_{O\lambda}$, $\hat{B}_{O\lambda}$, $C_{O\lambda}$, $\hat{C}_{O\lambda}$ aus den Koeffizienten der 1-D-Blöcke 1. oder 2. Grades berechnen lassen.

$$\text{Für } B_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}}) = \underline{z}_1^T B_{O\lambda} \underline{z}_2 \quad (5.3-86a)$$

$$C_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}}) = \underline{z}_1^T C_{O\lambda} \underline{z}_2 \quad (5.3-86b)$$

ergeben sich beim

Block 1.Grades mit der 1-D Teilübertragungsfunktion

$$H_{\lambda}(z) = \frac{b_{1\lambda} z + b_{0\lambda}}{z + c_{0\lambda}}$$

$b_{1\lambda}, b_{0\lambda}, c_{0\lambda}$ reell, die Vektoren

$$\underline{z}_1^T = [1 \quad z_1^{-\frac{1}{2}} \quad z_1^{-1}]^T \quad \underline{z}_2 = [z_2^{-\frac{1}{2}} \quad 1 \quad z_2^{-\frac{1}{2}}]^T$$

und die Koeffizientenmatrizen zu

$$\underline{B}_{0\lambda} = \begin{bmatrix} 0 & b_{0\lambda}^2 & 0 \\ -j \, b_{0\lambda} b_{1\lambda} & 0 & j \, b_{0\lambda} b_{1\lambda} \\ 0 & b_{1\lambda}^2 & 0 \end{bmatrix} \quad (5.3-87a)$$

$$\underline{C}_{0\lambda} = \begin{bmatrix} 0 & c_{0\lambda}^2 & 0 \\ -j \, c_{0\lambda} & 0 & j \, c_{0\lambda} \\ 0 & 1 & 0 \end{bmatrix} \quad (5.3-87b)$$

Für den Block 2.Grades mit der 1-D Teilübertragungsfunktion

$$H_{\lambda}(z) = \frac{b_{2\lambda} z^2 + b_{1\lambda} z + b_{0\lambda}}{z^2 + c_{1\lambda} z + c_{0\lambda}}$$

$b_{2\lambda}, b_{1\lambda}, b_{0\lambda}, c_{1\lambda}, c_{0\lambda}$ reell, gilt:

$$\underline{z}_1^T = [1 \quad z_1^{-\frac{1}{2}} \quad z_1^{-1} \quad z_1^{-\frac{3}{2}} \quad z_1^{-2}]^T$$

$$\underline{z}_2 = [z_2^{-1} \quad z_2^{-\frac{1}{2}} \quad 1 \quad z_2^{\frac{1}{2}} \quad z_2]^T$$

und

$$\underline{B}_{O\lambda} = \begin{bmatrix} 0 & 0 & b_{O\lambda}^2 & 0 & 0 \\ 0 & -j b_{1\lambda} b_{O\lambda} & 0 & j b_{1\lambda} b_{O\lambda} & 0 \\ -b_{O\lambda} b_{2\lambda} & 0 & b_{1\lambda}^2 & 0 & -b_{O\lambda} b_{2\lambda} \\ 0 & -j b_{1\lambda} b_{2\lambda} & 0 & j b_{1\lambda} b_{2\lambda} & 0 \\ 0 & 0 & b_{2\lambda}^2 & 0 & 0 \end{bmatrix}, \quad (5.3-88a)$$

$$\underline{C}_{O\lambda} = \begin{bmatrix} 0 & 0 & c_{O\lambda}^2 & 0 & 0 \\ 0 & -j c_{1\lambda} c_{O\lambda} & 0 & j c_{1\lambda} c_{O\lambda} & 0 \\ -c_{O\lambda} & 0 & c_{1\lambda}^2 & 0 & -c_{O\lambda} \\ 0 & -j c_{1\lambda} & 0 & j c_{1\lambda} c_{O\lambda} & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (5.3-88b)$$

Da

$$\hat{B}_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}}) = \underline{z}_1^T \underline{B}_{O\lambda}^* \underline{z}_2, \quad (5.3-89a)$$

$$\hat{C}_{O\lambda}(z_1^{\frac{1}{2}}, z_2^{\frac{1}{2}}) = \underline{z}_1^T \underline{C}_{O\lambda}^* \underline{z}_2 \quad (5.3-88b)$$

ist, sind damit die Teilübertragungsfunktionen $H_O(z_1, z_2)$, $\hat{H}_O(z_1, z_2)$ eindeutig bestimmt. Die spezielle Form der Koeffizientenmatrizen nach (5.3-87a,b) und (5.3-88a,b) bleibt auch für die Gesamtmatrizen $\underline{B}_O, \underline{C}_O$ der Teilübertragungsfunktionen

$$H_O(z_1, z_2) = \frac{\underline{z}_1^T \underline{B}_O \underline{z}_2}{\underline{z}_1^T \underline{C}_O \underline{z}_2} \quad (5.3-89)$$

$$\hat{H}_O(z_1, z_2) = \frac{\underline{z}_1^T \underline{B}_O^* \underline{z}_2}{\underline{z}_1^T \underline{C}_O^* \underline{z}_2}$$

erhalten. So sind die Koeffizienten der Matrizen

$$\begin{aligned} \underline{B}_0 &= \sum_{\lambda=1}^L \underline{B}_{0\lambda} = \underline{B}_{01} ** \underline{B}_{02} \cdot \cdot \cdot \underline{B}_{0L} \\ \underline{C}_0 &= \sum_{\lambda=1}^L \underline{C}_0 = \underline{C}_{01} ** \underline{C}_{02} \cdot \cdot \cdot \underline{C}_{0L} \end{aligned} \quad (5.3-90)$$

nur innerhalb des in (5.3-88a,b) gestrichelten Bereiches verschieden von Null. In diesem Bereich ist der Koeffizient für die Potenz

$$z_1^{\frac{\mu}{2}} z_2^{\frac{\nu}{2}} : \begin{cases} \text{rein reell,} & \text{falls } \mu, \nu \text{ gerade} \\ \text{rein imaginär,} & \text{falls } \mu \text{ und } \nu \text{ ungerade} \\ 0 & \text{sonst.} \end{cases}$$

Durch Summation ergibt sich

$$\begin{aligned} H_S(z_1, z_2) &= H_O(z_1, z_2) + \hat{H}_O(z_1, z_2) = \frac{B_S(z_1, z_2)}{C(z_1, z_2)} \\ &= \frac{\underline{z}_1^T \underline{B}_S \underline{z}_2}{\underline{z}_1^T \underline{C} \underline{z}_2} \end{aligned} \quad (5.3-91)$$

ein System mit rein reellen Koeffizienten.

Die Polynome $B_S(z_1, z_2), C(z_1, z_2)$ enthalten nur ganzzahlige Potenzen von z_1, z_2 , da

$$\begin{aligned} \underline{z}_1^T &= [1 \ z_1 \ \cdot \cdot \cdot \ z_1^{2n}] \\ \underline{z}_2 &= [z_2^{-n} \ \cdot \cdot \cdot 1 \ \cdot \cdot \cdot z_2^n] \end{aligned}$$

ist, wobei n den Filtergrad des 1-D-Prototypfilters angibt.

Das Nennerpolynom $C(z_1, z_2)$ setzt sich zusammen aus

$$C(z_1, z_2) = \prod_{\lambda=1}^{\Pi} C_{\lambda}(z_1, z_2) \cdot C_{\lambda}(z_1, z_2^{-1}), \quad (5.3-92)$$

wobei die Polynome C_{λ} für die Blöcke 1. und 2. Grades eindeutig durch die Koeffizienten des 1-D-Prototypfilters bestimmt sind und sich ergeben zu

$$C_{\lambda}(z_1, z_2) = \begin{cases} z_1 z_2 + c_{0\lambda}^2 & \text{Blockgrad 1} \\ z_1^2 z_2^2 + (c_{1\lambda}^2 - 2c_{0\lambda}) z_1 z_2 + c_{0\lambda}^2 & \text{Blockgrad 2} \end{cases} \quad (5.3-93)$$

Aus Gl. (5.3-92) und (5.3-93) ist erkennbar, daß das rekursive System

$$H_r(z_1, z_2) = \frac{1}{C(z_1, z_2)}$$

als Kaskade eines um $\pm \frac{\pi}{4}$ gedrehten 1-D-Teilfilters (Kap. 5.3.2)

$$H_r(z) = \frac{1}{C(z)} \quad (5.3-94)$$

mit
$$C(z) = \prod_{\lambda=1}^{\Pi} C_{\lambda}(z)$$

und

$$C_{\lambda}(z) = \begin{cases} z + c_{0\lambda}^2 & \text{Blockgrad 1} \\ z^2 + (c_{1\lambda}^2 - 2c_{0\lambda})z + c_{0\lambda}^2 & \text{Blockgrad 2} \end{cases} \quad (5.3-95)$$

gedeutet werden kann.

Die Nennerkoeffizientenmatrix $\underline{B}_S$ in (5.3-91) berechnet sich zu

$$\underline{B}_S = 2 \cdot \operatorname{Re} \left\{ \prod_{\lambda=1}^{\Pi} \underline{B}_{0\lambda} \right\} \cdot \underline{C}_{0\lambda}^* \quad (5.3-96)$$

wobei alle Koeffizienten für die nichtganzzahligen Potenzen von z_1 und z_2 zu streichen sind.

Die Koeffizientenmatrix $\underline{B}_S$ hat wie die Matrizen $\underline{B}_{O\lambda}, \underline{C}_{O\lambda}$ (5.3-87a) bis (5.3-88b) die Eigenschaft, daß die Koeffizienten außerhalb des in Bild 5.15 dargestellten Bereiches Null werden

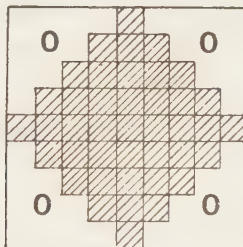


Bild 5.15: Bereich, in dem die Koeffizienten der Matrix $\underline{B}_S$ von Null verschieden sein können (schraffiert).

Durch die Summenbildung nach (5.3-91) entstehen für $\Omega_1, \Omega_2 \in \{0, \pi\}$ Überlappungen der Durchlaßbereiche bzw. Übergangsbereiche der Teilfilter $H_O(z_1, z_2), \hat{H}_O(z_1, z_2)$. Durch Verkleinerung der Durchlaßgrenze oder durch Nachoptimierung des Zählerpolynoms kann dieser Effekt reduziert werden.

Das Entwurfsverfahren hat folgende Vorteile:

- Das resultierende zweidimensionale Filter ist stabil, sofern das 1-D-Prototypfilter stabil ist.
- Für die Realisierung kann das Filter in eine Kaskade eines nichtrekursiven Systems und eines rekursiven Allpolfilters mit separablem Nennerpolynom zerlegt werden. Dazu wird das um 45° gedrehte Eingangsbild zunächst mit einem nichtrekursiven Filter und anschließend zeilen-

und spaltenweise mit einem 1-D rekursiven System gefiltert. Für linearphasige Filter wird das nichtre-kursive System als linearphasiges 2-D-Filter realisiert. Die zeilen- und spaltenweise 1-D rekursive Filterung kann als 1-D Nullphasenfilterung [59] exakt durchgeführt werden.

- Durch den Entwurf des 1-D-Prototypfilters nach einem vorgegebenen Toleranzschema liegt der Grad des zweidimensionalen Gesamtsystems fest, d.h. die Unsicherheit bei der Wahl des Grades für Zähler- und Nennerpolynom entfällt.

Ein Beispiel soll die Brauchbarkeit dieses Verfahrens verdeutlichen.

5.3.3.2 Beispiel

Das 1-D-Prototypfilter wurde wieder mit dem Programm DORE-DI [82] entworfen.

Für $\Omega_D=0.5\pi$, $\Omega_S=0.7\pi$ ergibt sich ein Cauerfilter 4.Grades mit $\delta_D<1\%$, $\delta_S<1\%$ und den Koeffizienten

KONSTANTE = 0.1495069E+00

B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1261904E+01	C1 =	0.2317095E+00
B0 =	0.1000000E+01	C0 =	0.6809576E+00
B2 =	0.1000000E+01	C2 =	0.1000000E+01
B1 =	0.1834764E+01	C1 =	-0.1365015E+00
B0 =	0.1000000E+01	C0 =	0.1239722E+00

Zur Berechnung der Koeffizienten des Fan-Filters wurde ein FORTRAN-Programm an einer PDP 11/45 erstellt.

Das Nennerpolynom $C(z_1, z_2)$ für $H_S(z_1, z_2)$ (5.3-91) ergibt sich zu

$$C(z_1, z_2) = C_1(z_1 \cdot z_2) \cdot C_1(z_1 \cdot z_2^{-1}) ,$$

wobei nach (5.3-95)

$$C_1(z) = (z^2 - 1.3082260 \cdot z + 0.463703) \\ \cdot (z^2 - 0.2293118 \cdot z + 0.0153691)$$

wird.

In Bild 5.16a ist der Betragsfrequenzgang des resultierenden Fan-Filters angegeben.

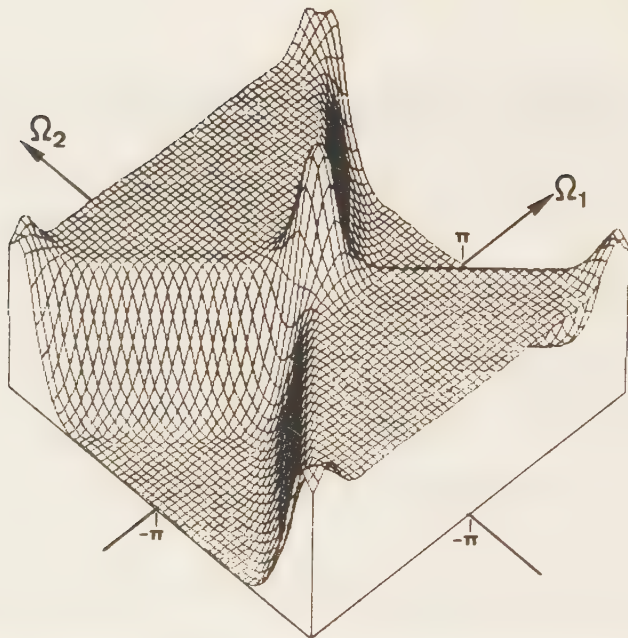
Eine Nachoptimierung des Zählerpolynoms führt zu dem Betragsfrequenzgang nach Bild 5.16b. Die zugehörigen Zählerkoeffizientenmatrizen sind in Tabelle 5.1 angegeben.

5.3.4 Entwurfsverfahren 3

Bei vielen Anwendungen der Bildverarbeitung sind zur Vorfilterung Glättungsfilter notwendig. Dabei kommt es im allgemeinen nicht so sehr auf die exakte Einhaltung eines vorgeschriebenen Toleranzschemas, sondern vielmehr auf die schnelle Durchführbarkeit der Vorfilterung an. In diesem Sinne ist gerade das Filter optimal, welches den gewünschten Glättungseffekt bewirkt und zudem möglichst geringen Rechenaufwand erfordert (siehe Kap.6.1). Das gewünschte Filter soll also niedergradig und die Koeffizienten möglichst einfach (z.B. Zweierpotenzen) sein.

Zur Erfüllung dieser Forderungen wurde die nachfolgend beschriebene Methode angewandt. Der Vorteil dieses Verfahrens liegt darin, daß sich die Koeffizienten des 2-D-Filters analytisch aus den Koeffizienten des 1-D-Filters berechnen lassen.

a)



b)

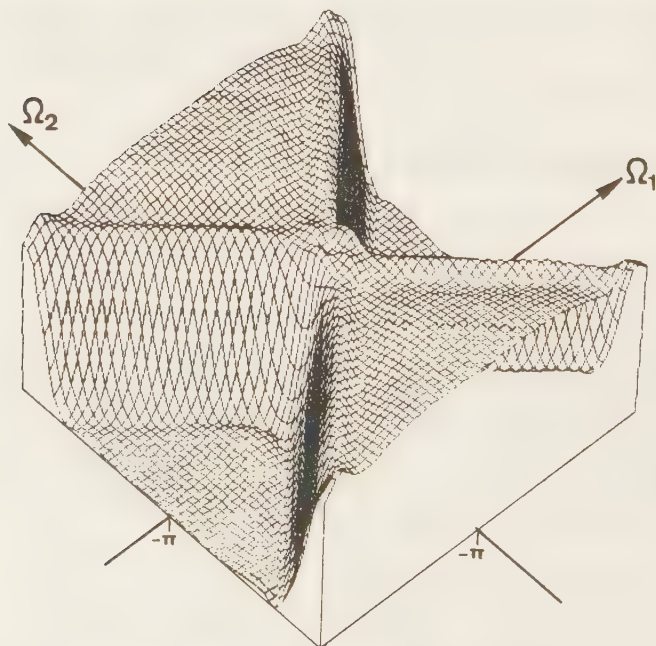


Bild 5.16: Betragsfrequenzgang des Fan-Filters nach 5.3.3.2

[illegible]

b) Zählerkoeffizientenmatrix des Filters nach Bild 5.15b

[illegible]

gemeinsamer konstanter Vorfaktor: 0.02458754

Tabelle 5.1: Zählerkoeffizientenmatrizen der Fan-Filter nach Bild 5.15

5.3.4.1 Prinzip

Ausgehend von einem 1-D-Prototypfilter mit der Übertragungsfunktion (5.3-3)

$$H_1(z) = \prod_{\lambda=1}^1 H_{1\lambda}(z)$$

wird jedes der 1-D Teilfilter 1. oder 2. Grades mit der Teilübertragungsfunktion $H_{1\lambda}(z)$ durch

$$z \rightarrow g(z_1, z_2) = (1 - \gamma_\lambda) z_1 \cdot z_2 + \gamma_\lambda (z_1 + z_2) - \gamma_\lambda \quad (5.3-97)$$

in ein 2-D-Teilblock $H_{2\lambda}(z_1, z_2)$ transformiert, wodurch ein 2-D System mit der Übertragungsfunktion

$$H_2(z_1, z_2) = \prod_{\lambda=1}^1 H_{2\lambda}(z_1, z_2) \quad (5.3-98)$$

entsteht. Dabei kann der Transformationsparameter γ_λ für Zähler- und Nennerpolynom jedes Blockes verschiedene Werte annehmen. Blöcke 2. Grades werden in zwei Blöcke 1. Grades mit konjugiert komplexen Koeffizienten zerlegt und mit γ_λ bzw. γ_λ^* transformiert und danach wieder zu einem Block mit reellen Koeffizienten zusammengefaßt.

So ergibt sich aus einem eindimensionalen Block 1. Grades mit der Übertragungsfunktion

$$H_{1\lambda}(z) = \hat{H}(z) = \frac{b_1 z + b_0}{z + c_0} \quad (5.3-99)$$

nach der Transformation

$$H_{2\lambda}(z_1, z_2) = \frac{[1 \ z_1] \begin{bmatrix} b_0 - \gamma_1 b_1 & b_1 \gamma_1 \\ b_1 \gamma_1 & b_1 (1 - \gamma_1) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ z_2 \end{bmatrix}}{[1 \ z_1] \begin{bmatrix} c_0 - \gamma_2 & \gamma_2 \\ \gamma_2 & 1 - \gamma_2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ z_2 \end{bmatrix}}, \quad (5.3-100)$$

wobei γ_1, γ_2 die reellen Transformationsparameter für Zähler und Nennerpolynom bezeichnen.

Für den Block 2.Grades mit der Übertragungsfunktion

$$H_{1\lambda}(z) = H(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^2 + c_1 z + c_0} =$$

$$= b_2 \cdot \frac{(z - z_0)(z - z_0^*)}{(z - z_\infty)(z - z_\infty^*)}, \quad (5.3-101)$$

wobei $b_1 = b_2(z_0 + z_0^*) = -2b_2 \cdot \operatorname{Re}\{z_0\}$

$$b_0 = b_2 \cdot z_0 \cdot z_0^* = b_2 |z_0|^2 \quad (5.3-102)$$

$$c_1 = -(z_\infty + z_\infty^*) = -2 \cdot \operatorname{Re}\{z_\infty\}$$

$$c_0 = z_\infty \cdot z_\infty^* = |z_\infty|^2$$

ist, ergibt sich durch Anwendung der Transformation nach (5.3-97) mit

$$\gamma_1 = \lambda_1 + j \cdot \rho_1 \quad (5.3-103)$$

$$\gamma_2 = \lambda_2 + j \cdot \rho_2$$

auf den 1-D Teilblock in (5.3-101) mit den komplexen Koeffizienten z_0, z_∞ und mit γ_1^*, γ_2^* auf den 1-D Teilblock mit z_0^*, z_∞^*

der zweidimensionale Block 2.Grades zu

$$H_{2\lambda}(z_1, z_2) = \frac{\underline{z}_1^T \underline{B} \underline{z}_2}{\underline{z}_1^T \underline{C} \underline{z}_2} \quad (5.3-104)$$

$$\text{mit } \underline{z}_1^T = [1, z_1^{-1}, z_1^{-2}]$$

$$\underline{z}_2 = [1, z_2^{-1}, z_2^{-2}]^T$$

$$\underline{B} = b_2 \cdot \begin{bmatrix} \hat{b}_{00} & \hat{b}_{01} & \hat{b}_{02} \\ \hat{b}_{10} & \hat{b}_{11} & \hat{b}_{12} \\ \hat{b}_{20} & \hat{b}_{21} & \hat{b}_{22} \end{bmatrix}, \quad \underline{C} = \begin{bmatrix} c_{00} & c_{01} & c_{02} \\ c_{10} & c_{11} & c_{12} \\ c_{20} & c_{21} & c_{22} \end{bmatrix}, \quad (5.3-105)$$

wobei

$$\begin{aligned} \hat{b}_{00} &= 1 - 2 \operatorname{Re}\{\gamma_1\} + |\gamma_1|^2 \\ \hat{b}_{01} &= \hat{b}_{10} = 2(\operatorname{Re}\{\gamma_1\} - |\gamma_1|^2) \\ \hat{b}_{02} &= \hat{b}_{20} = |\gamma_1|^2 \\ \hat{b}_{11} &= 2(\operatorname{Re}\{\gamma_1 z_0^*\} + 2|\gamma_1|^2 - \operatorname{Re}\{\gamma_1\} - \operatorname{Re}\{z_0\}) \\ \hat{b}_{12} &= \hat{b}_{21} = -2(\operatorname{Re}\{\gamma_1 z_0^*\} + |\gamma_1|^2) \\ \hat{b}_{22} &= |\gamma_1|^2 + 2 \operatorname{Re}\{\gamma_1 z_0^*\} + |z_0|^2 \quad \text{gilt.} \end{aligned} \quad (5.3-106)$$

Die Koeffizienten der Matrix $\underline{C}$ lassen sich aus (5.3-106) bestimmen, wenn $\hat{b}_{\mu\nu}$ durch $c_{\mu\nu}$, γ_1 durch γ_2 und z_0 durch z_∞ ersetzt wird.

Für den Fall

$$\begin{aligned} \gamma_1 &= \frac{-z_0}{1-z_0} \\ \gamma_2 &= \frac{-z_\infty}{1-z_\infty} \end{aligned} \quad (5.3-107)$$

ist das transformierte System separabel, d.h.

$$H_{2\lambda}(z_1, z_2) = H_0(z_1) \cdot H_0(z_2).$$

Für den Sonderfall

$$\begin{aligned} \gamma_1 &= -z_0 \\ \gamma_2 &= -z_\infty \end{aligned} \quad (5.3-108)$$

ist $\hat{b}_{22}=\hat{b}_{12}=\hat{b}_{21}=c_{22}=c_{12}=c_{21}=0$, die Koeffizientenmatrizen $\underline{B}, \underline{C}$ haben deshalb die Dreiecksform:

$$\underline{B} = \begin{bmatrix} b_2 - b_1 + b_0 & -(2b_0 - b_1) & b_0 \\ -(2b_0 - b_1) & 2b_0 & 0 \\ b_0 & 0 & 0 \end{bmatrix} \quad (5.3-109a)$$

$$\text{und} \quad \underline{C} = \begin{bmatrix} 1 - c_1 + c_0 & -(2c_0 - c_1) & c_0 \\ -(2c_0 - c_1) & 2c_0 & 0 \\ c_0 & 0 & 0 \end{bmatrix} \quad (5.3-109b)$$

Die Untersuchung, für welches γ_2 die Transformation nach (5.3-97) auf stabile Filter führt, wird am einfachsten wieder im w -Bereich durchgeführt.

Die Transformation im z -Bereich nach (5.3-97) ist äquivalent mit der Transformation

$$w = \frac{-2\gamma w_1 w_2 + w_1 + w_2}{(1-2\gamma) w_1 w_2 + 1} \quad (5.3-110)$$

im w -Bereich.

Durch bilineare Transformation (5.3-5) der Teilübertragungsfunktion

$$\tilde{H}(z) = \frac{1}{z - z_\infty} \quad (5.3-111)$$

in den w -Bereich ergibt sich die Forderung, daß

$$N(w_1, w_2) = 1 + \left(1 - \frac{4\gamma_2}{1 - z_\infty}\right) w_1 w_2 + \frac{1 + z_\infty}{1 - z_\infty} (w_1 + w_2) \quad (5.3-112)$$

ein stabiles Polynom (5.3-33a,b) sein muß.

Ein Vergleich mit (5.3-30) zeigt, daß für $\gamma_2 = \lambda_2 + j\rho_2$

$$\left. \begin{aligned} \text{und } \alpha &= 1 + \frac{2}{d_0} [\rho_2 d_3 - \lambda_2 (2+c_1)] \\ \beta &= [\rho_2 (2+c_1) + \lambda_2 d_3] \cdot \left(-\frac{2}{d_0}\right) \end{aligned} \right\} \quad (5.3-113)$$

mit d_0, d_3 nach (5.3-27) und $c_1 = -2 \operatorname{Re}\{z_\infty\}$ beide Beziehungen identisch sind.

Für reelle Pole ist das resultierende 2-D-Filter stets stabil, sofern

$$\gamma_2 = \lambda_2 < \frac{1-z_\infty}{4} \quad (5.3-114)$$

ist. Für reelle Pole zwischen

$$-0.4 < z_\infty < 1.0$$

ist die mittlere quadratische Abweichung des Betragsfrequenzgangs jedes Blocks (5.3-111) vom idealen zirkularsymmetrischen Verlauf minimal, wenn

$$\gamma_2 = \lambda_2 = -\frac{z_\infty}{2}$$

gewählt wird.

Für komplexe Pole lassen sich aus (5.3-112) durch Vergleich mit (5.3-30) die Transformationsparameter ρ_2, λ_2 als Funktionen von α und β berechnen zu

$$\lambda_2 = \frac{1}{8} [(2+c_1)(1-\alpha) + d_3\beta] , \quad (5.3-115a)$$

$$\rho_2 = \frac{1}{8} [d_3(1-\alpha) - (2+c_1)\beta] . \quad (5.3-115b)$$

Wird α und β nach (5.3-43) bzw. (5.3-44) gewählt, ergibt sich stets ein stabiles resultierendes 2-D-Filter, sofern das 1-D-Prototypfilter stabil ist.

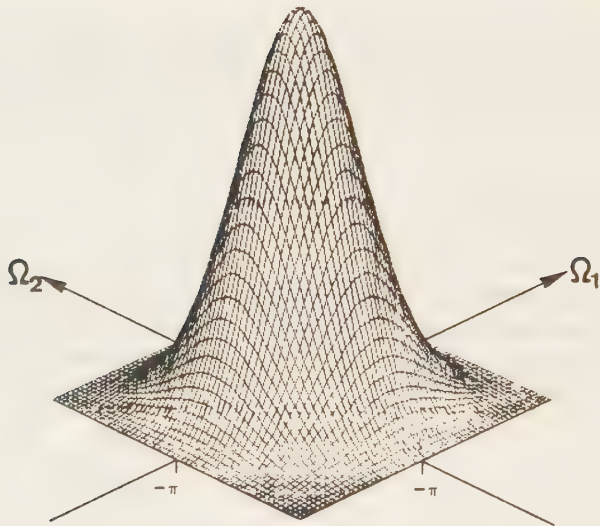
5.3.4.2 Beispiel

Ausgehend von einem 1-D-Prototyptiefpfaß mit einer doppelten Nullstelle bei $z=-1$ und einer doppelten reellen Polstelle bei $z=\frac{1}{7}$ ergibt die Transformation nach (5.3-97) ein zweidimensionales Filter mit der Übertragungsfunktion

$$H(z_1, z_2) = \frac{9}{4} \cdot \frac{[1, z_1^{-1}, z_1^{-2}] \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ z_2^{-1} \\ z_2^{-2} \end{bmatrix}}{[1, z_1^{-1}, z_1^{-2}] \begin{bmatrix} 64 & -16 & 1 \\ -16 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ z_2^{-1} \\ z_2^{-2} \end{bmatrix}}, \quad (5.3-116)$$

wenn das Zählerpolynom separabel (5.3-20) und das Nennerpolynom mit $\gamma_2 = -z_\infty$ (5.3-108) transformiert werden. Dieses Filter hat neben der positiven Eigenschaft einfacher Koeffizienten den Vorzug, daß der Betragsfrequenzgang näherungsweise zirkularsymmetrisch (Bild 5.17) ist. Das Filter besitzt außerdem näherungsweise lineare Phase, weshalb die Filterung in einer Rekursionsrichtung ausreicht, um Phasenverzerrungen im Ausgangsbild $|5|$ vernachlässigbar klein zu halten.

Der Entwurf eines derartigen einfachen Filters wurde durch den in 6.1 beschriebenen Anwendungsfall angeregt. Es zeigt sich jedoch, daß es auch für andere Anwendungen ($|94|$, Kap.6.2) erfolgreich eingesetzt werden kann.



Höhenlinienabstand:
0.1, 0.2, ..., 0.9

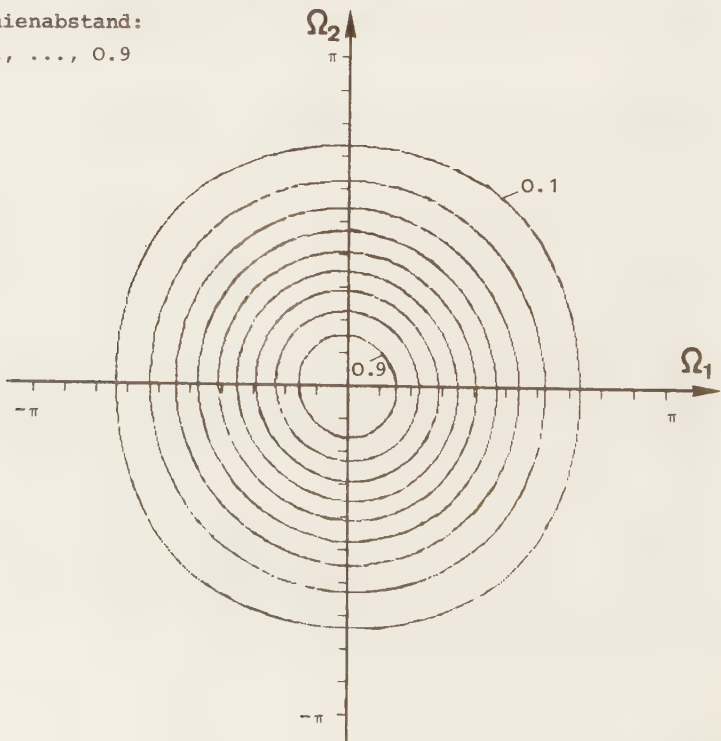


Bild 5.17: Betragsfrequenzgang des Filters nach (5.3-116)

Im nachfolgenden Kapitel werden einige Anwendungen rekursiver digitaler Filter für typische Aufgabenstellungen der zweidimensionalen Signalverarbeitung aufgezeigt.

6. Anwendungsbeispiele für 2-D rekursive digitale Filter

Einige typische und häufig wiederkehrende Aufgabenstellungen bei der Verarbeitung zweidimensionaler digitaler Signale sind

- Glättung bzw. Rauschunterdrückung
- Kantendetektion bzw. Konturextraktion
- Hervorhebung bestimmter Bilddetails.

Im folgenden wird gezeigt, wie zur Lösung vorteilhaft 2-D rekursive digitale Filter eingesetzt werden können.

6.1 Rauschunterdrückung bei der automatischen Auswertung von Mikro-Chromatogrammen *)

Zum Nachweis und der quantitativen Bestimmung von Aminosäuren in organischen Flüssigkeiten wird das Verfahren der zweidimensionalen Dünnschicht-Chromatographie angewendet [95-97]. Dazu werden die Aminosäuren in Verbindungen mit Dansylchlorid übergeführt, die bei UV-Bestrahlung fluoreszieren und damit sichtbar werden.

Für die Chromatographie wird das Substanzgemisch in geringer Menge auf eine Trägerfolie (3x4cm) in der linken unteren Ecke aufgetragen und getrocknet. Durch Lösung mit geeigneten Laufmitteln erreicht man eine Trennung der unterschiedlichen Aminosäuren. Unter Bestrahlung mit UV-Licht ergibt

*) Die hier beschriebene Aufgabenstellung ergab sich aus der Zusammenarbeit mit dem Max-Planck-Institut für experimentelle Medizin, Forschungsstelle Neurochemie (Prof. Dr. V. Neuhoff) in Göttingen.

sich ein Bild mit verschiedenen starken Flecken. Dabei gibt die Lage eines Flecks innerhalb des Bildes Aufschluß über die Art der Aminosäure, die Intensität und die Fleckausdehnung ist ein Maß für die Menge der vorhandenen Aminosäure. Ein Fleck entspricht etwa 10^{-10} bis 10^{-12} Mol der Aminosäure. Die Digitalisierung erfolgt durch ein Auflicht-Skanning-Mikroskop-Photometer und liefert ein Bild mit 300×400 Bildpunkten und 4096 Graustufen [95]. Um den notwendigen Aufwand bei der automatischen Fleckerkennung und Flecktrennung an einer PDP-8 zu verringern, muß das 2-D Signal geglättet werden. Für diese Aufgabe wurde das rekursive Filter nach 5.3.4.2 entworfen. Die Anforderungen dabei sind

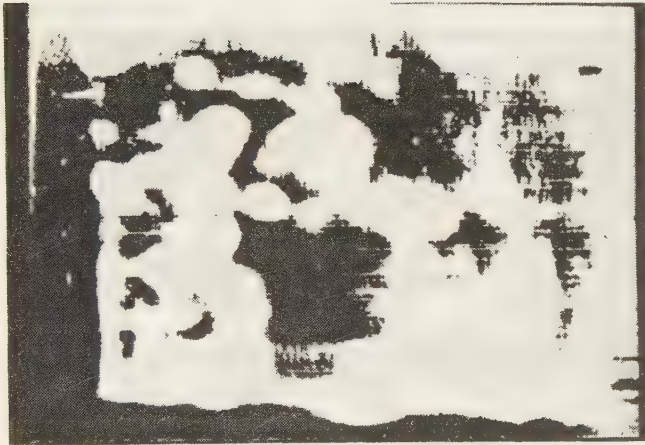
- geringer Rechenaufwand für die Filterung,
- keine Verfälschung der Signalamplituden.

Es zeigt sich, daß es bei diesen Signalen Vorteile bringt, die Filterung auf ein durch $\sqrt{I}$ kommandiertes Bild anzuwenden, wobei I die Intensität des jeweiligen Bildpunkts angibt [94].

In Bild 6.1a ist eine Intensitätsebene (574)₈ eines ungeglätteten Chromatogramms dargestellt. Bild 6.1b zeigt die gleiche Ebene nach der rekursiven Filterung.

Die Wirkung auf die Fleckerkennung verdeutlicht Bild 6.2. Während bei dem ungeglätteten Signal rauschartige Störungen, hervorgerufen durch Schwankungen der Lampenintensität, Inhomogenität des Trägermaterials, Staubpartikelchen etc., als kleine Flecke erkannt werden (Bild 6.2a), wird dies durch die Anwendung des rekursiven Vorfilters weitgehend vermieden (Bild 6.2b).*)

*) Die Photographien wurden freundlicherweise von Herrn Dr. H.-G. Zimmer MPI, Göttingen zur Verfügung gestellt.

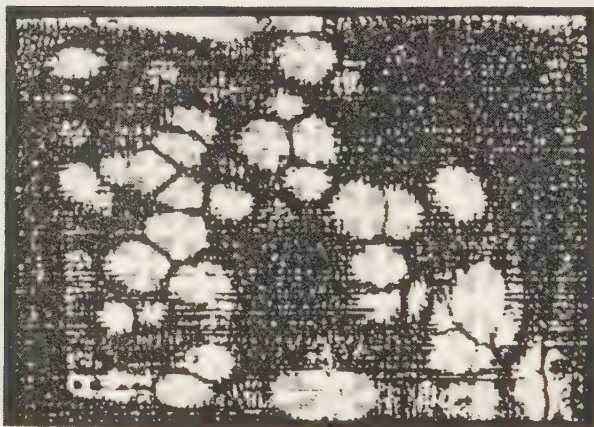


a) Original

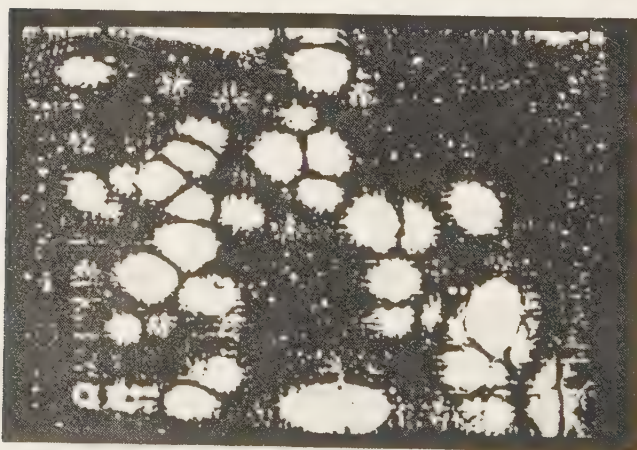


b) mit dem rek. Tiefpaß nach (5.3-116) geglättet

Bild 6.1: Bitebene eines Chromatogramms vor und nach der Glättung



a) Original



b) geglättet

Bild 6.2: Auswirkung der Vorfilterung auf die Fleckerkennung

6.2 Rekursive Hochpaßfilterung zur Hervorhebung von Bilddetails

Die Computertomographie (CT)-Aufnahme nach Bild 6.3a zeigt einen Querschnitt durch den Brustkorb eines Menschen. Dabei interessieren besonders die Verästelungen der Bronchialwege. Zur Hervorhebung dieser Bilddetails wird eine rekursive Hochpaßfilterung angewendet. Als Hochpaßfilter wird das Filter

$$\hat{H}(z_1, z_2) = 1 - H(z_1, z_2) \cdot H(z_1^{-1}, z_2^{-1})$$

verwendet, wobei $H(z_1, z_2)$ die Übertragungsfunktion des Tiefpasses nach Gl. (5.3-116) beschreibt.

Wie aus Bild 6.3b zu sehen ist, wird damit die gewünschte Hervorhebung der Bronchialwege erreicht. Noch deutlicher wird dies im Negativ (Bild 6.3c).

6.3 Anwendung eines rekursiven Kantendetektionsfilters

Operatoren zur Kantendetektion wurden in der Diplomarbeit [98] ausführlich untersucht. Ausgehend von einem statistischen Kantenmodell nach [87] läßt sich ein bzgl. des mittleren quadratischen Fehlers optimales Laplacefilter mit gaußscher Bandbegrenzung berechnen.

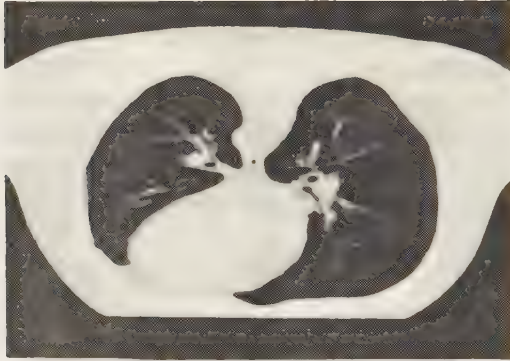
Die Übertragungsfunktion des kontinuierlichen Systems ergibt sich zu

$$H(\omega) = \frac{\omega^2 e^{-\omega^2/2}}{1 + \frac{[\omega^2 + 2(1-\rho)^2 \lambda^2] \sqrt{\omega^2 + (1-\rho)^2 \lambda^2}}{8(1-\rho) \lambda \xi}} .$$

Dabei bedeutet

$\omega = 2\pi f$ die Kreisfrequenz ,

ρ den Korellationsgrad einer Kante ($|\rho| < 1$),



a) Original



b) Hochpaßfilterung (positiv)



c) Hochpaßfilterung (negativ)

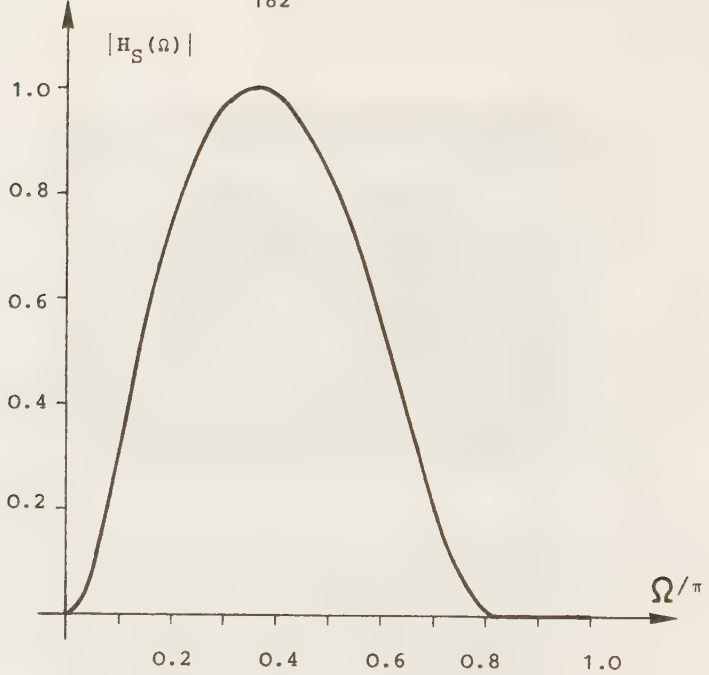
- λ ein Maß für die mittlere Zahl der Kanten pro Längeneinheit,
- ξ das Signal-Rauschverhältnis der Kantenstruktur gegenüber Rauschen und anderen Bilddetails.

Durch geeignete Wahl der Parameter ρ, λ, ξ kann eine Anpassung an spezielle Bilder vorgenommen werden.

Für $\lambda=0.125$, $\xi=3\text{dB}$, $\rho=0.5$ erhält man im z -Bereich durch bilineare Transformation (5.3-15) den in Bild 6.4a dargestellten Sollfrequenzgang $|H_S(\Omega)|$. Approximiert wird die Funktion $\sqrt[4]{|H_S(\Omega)|}$ durch ein eindimensionales rekursives Filter 2.Grades (Bild 6.4b). Der Entwurf erfolgte mit einem Programm von Deczky [20,80]. Nach der in Kap.5.1 beschriebenen Methode läßt sich aus diesem 1-D-Filter ein zweidimensionales rekursives Filter entwerfen. In Bild 6.5 ist der Betragsfrequenzgang dargestellt, der sich ergibt, wenn 4 derartige Filter mit jeweils verschiedener Rekursionsrichtung (4.3-3a) kaskadiert werden.

Bild 6.6b zeigt die Wirkung dieses Filters auf ein CT-Bild, das einen Schnitt durch die Wirbelsäule eines Menschen darstellt (Bild 6.6a). Bild 6.6c zeigt das Ergebnis bei Anwendung des Hochpaßfilters von Kap. 6.2.

a)



b)

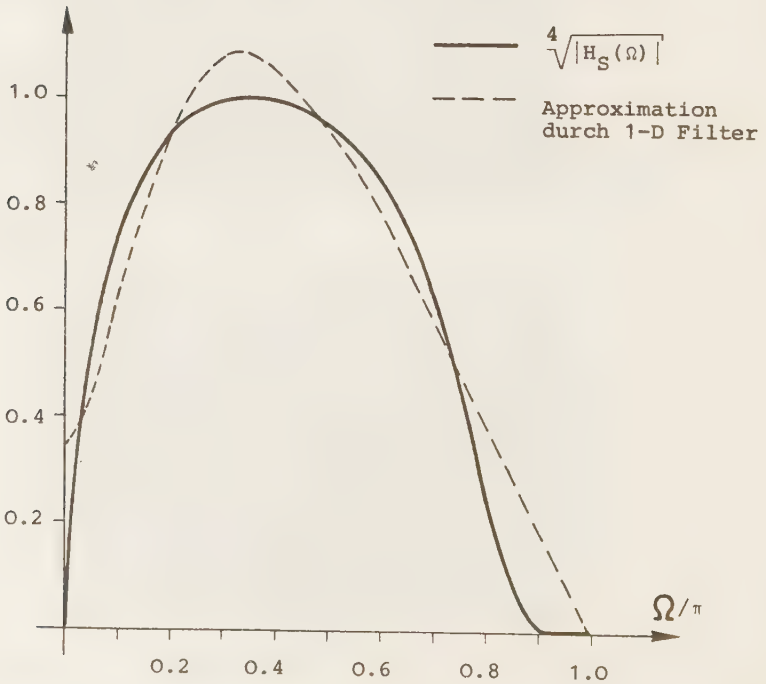


Bild 6.4: Sollfrequenzgang des 1-D Kantendetektionsfilters

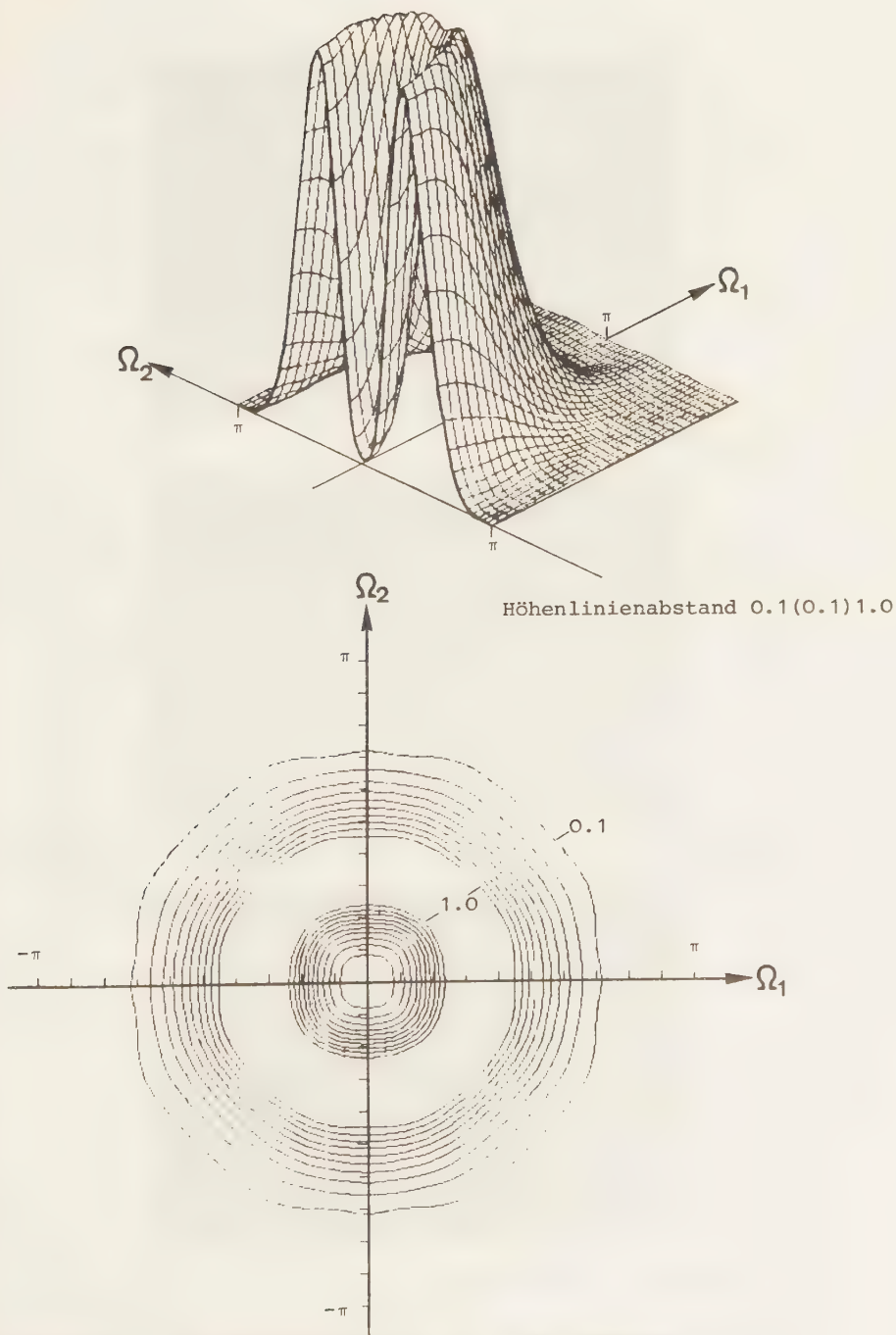
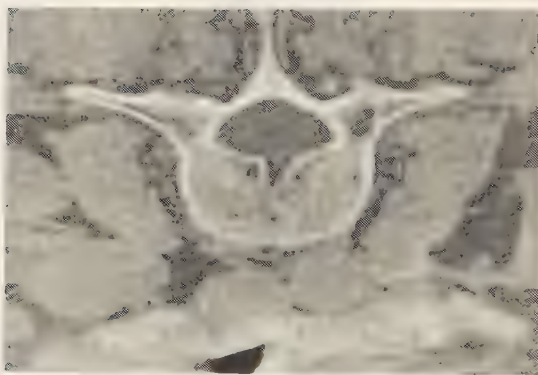


Bild 6.5: Betragsfrequenzgang des 2-D Laplacefilters mit gaußscher Bandbegrenzung



a) Original



b) nach der Filterung mit dem Kantendetektionsfilter (Bild 6.5)



c) hochpaßgefiltertes Signal

7. Zusammenfassung

In der vorliegenden Arbeit wurden folgende Probleme der zweidimensionalen digitalen Signalverarbeitung behandelt:

- Randeffekte, ortsvariante Filterung,
- Stabilitätstests für rekursive 2-D-Filter,
- Realisierungsaufwand für die Durchführung der digitalen Filterung an einem Minicomputer,
- Entwurf stabiler rekursiver zweidimensionaler Filter.

Zunächst wurde die Äquivalenz der ortsvarianten nichtrekursiven Filterung mit der ortsinvarianten Filterung eines am Rand extrapolierten Eingangsbildes zur Verringerung der störenden Randeffekte aufgezeigt. Eine genaue Analyse des Realisierungsaufwands für die Durchführung der nichtrekursiven und rekursiven Filterung an einem Minicomputer PDP 11/45 führt zu Aussagen darüber, wann welche Art der Filterung vorteilhaft ist.

Ein neuer numerischer Stabilitätstest und neue Entwurfsverfahren, die stabile rekursive Filter garantieren, also einen derartigen Test überflüssig machen, bilden einen Beitrag zur Theorie und der Anwendbarkeit rekursiver zweidimensionaler Systeme.

Anhang 1:Herleitung der Beziehungen (4.2-3a,b)

$$\text{Mit } C(z_1, z_2) = [1, z_1^{-1}] \cdot \begin{bmatrix} 1 & c_{01} \\ c_{10} & c_{11} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ z_2^{-1} \end{bmatrix}$$

$c_{01}, c_{10}, c_{11} \in \mathbb{C}$ ergibt sich

$$A(z_1, z_2) = z_1 z_2 C(z_1, z_2) \text{ zu}$$

$$A(z_1, z_2) = [1, z_1] \cdot \begin{bmatrix} c_{11} & c_{10} \\ c_{01} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ z_2 \end{bmatrix}.$$

Dieses zweidimensionale Polynom in z_1, z_2 ist darstellbar als

$$A(z_1, z_2) = A_1(z_1) z_2 + A_0(z_1) = F(z_2) \quad (\text{A1-1})$$

$$\text{mit } A_0(z_1) = c_{01} z_1 + c_{11} \quad (\text{A1-2})$$

$$A_1(z_1) = z_1 + c_{10} \quad (\text{A1-3})$$

Als notwendige Bedingung für die Stabilität muß

$$A_1(z_1) \neq 0 \quad \text{für } |z_1| \geq 1 \quad (\text{A1-4})$$

sein, daraus folgt nach (A1-3)

$$|c_{10}| < 1,$$

was der Ungleichung (4.2-3a) entspricht.

Zum Nachweis der Bed. (4.2-3b) wird die Beziehung (A1-1) betrachtet. Für das Polynom $F(z_2)$ mit den komplexen

Koeffizienten $A_0(z_1), A_1(z_1)$ lassen sich mit der Marden-Jury-Tafel [43, 56] die folgenden notwendigen und hinreichenden Bedingungen angeben. Es muß

$$B_0(z_1) = \begin{bmatrix} A_0(z_1) & A_1(z_1) \\ A_1^*(z_1) & A_0^*(z_1) \end{bmatrix} = |A_0(z_1)|^2 - |A_1(z_1)|^2 < 0 \quad (A1-5)$$

für $|z_1|=1$ sein.

Außerdem muß (A1-4) erfüllt werden.

$$\text{Mit } |A_0(z_1)|^2 = |c_{01}|^2 |z_1|^2 + |c_{11}|^2 + c_{01} c_{11}^* z_1 + c_{01}^* c_{11} z_1^*$$

$$\text{und } |A_1(z_1)|^2 = |z_1|^2 + c_{10} z_1^* + c_{10}^* z_1 + |c_{10}|^2$$

ergibt sich

$$B_0(z_1) = |c_{11}|^2 - |c_{10}|^2 + |c_{01}|^2 |z_1|^2 + (c_{01} c_{11}^* - c_{10}^*) z_1 + (c_{01}^* c_{11} - c_{10}) z_1^* - |z_1|^2.$$

Für $|z_1|=1$ ergibt sich mit $z_1 = e^{j\Omega_1}$

$$\begin{aligned} B_0(z_1) \Big|_{|z_1|=1} &= \tilde{B}_0(\Omega_1) = \\ &= |c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1 \\ &\quad + 2 \operatorname{Re}\{(c_{01} c_{11}^* - c_{10}^*) \cdot e^{j\Omega_1}\}. \end{aligned} \quad (A1-6)$$

Mit der Substitution

$$r = |c_{01} c_{11}^* - c_{10}^*| = |c_{01}^* c_{11} - c_{10}| \quad (A1-7)$$

$$\varphi = \angle (c_{01} c_{11}^* - c_{10}^*)$$

ergibt sich aus (A1-6)

$$\begin{aligned}\tilde{B}_O(\Omega_1) &= |c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1 \\ &+ 2r \cdot \cos(\Omega_1 + \varphi).\end{aligned}\tag{A1-8}$$

Für $x = \cos(\Omega_1 + \varphi)$

ergibt sich aus (A1-8) und (A1-5)

$$\begin{aligned}\hat{B}_O(x) &= |c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1 + 2rx < 0 \\ \text{für } \forall x \text{ mit } |x| \leq 1\end{aligned}\tag{A1-9}$$

Die zugehörige Gleichung zu (A1-9) stellt eine Geraden-
gleichung dar. Deshalb ist (A1-9) stets erfüllt, wenn

$$\begin{aligned}B_O(-1) &= |c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1 - 2r < 0 \\ \text{und } \hat{B}_O(1) &= |c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1 + 2r < 0 \\ \text{gilt. Da } \hat{B}_O(0) &= |c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1,\end{aligned}\tag{A1-10}$$

ergibt sich

$$\hat{B}_O(-1) = \hat{B}_O(0) - 2r < 0\tag{A1-11}$$

$$\hat{B}_O(1) = \hat{B}_O(0) + 2r < 0.\tag{A1-12}$$

Da nach (A1-7) $r \geq 0$ ist, ist Ungl. (A1-11) automatisch
erfüllt, falls Ungl. (A1-12) erfüllt wird.

Mit (A1-7), (A1-10) folgt dann aus (A1-12)

$$|c_{11}|^2 + |c_{01}|^2 - |c_{10}|^2 - 1 + 2|c_{01}^* c_{11} - c_{10}| < 0,$$

was Bedingung (4.2-3b) darstellt.

Anhang 2:

Nachweis, daß die Transformationsparameter

$$\alpha = -\left(\frac{d_3}{d_0}\right)^2 + \varepsilon + x^2, \quad x \in \mathbb{R} \quad (5.3-40)$$

$$\beta = \beta_{\min} + \varepsilon + (\Delta\beta - 2\varepsilon) \sin^2\phi \quad (5.3-41)$$

mit

$$\Delta\beta = \beta_{\max} - \beta_{\min} = 2 \frac{d_1}{d_0} \sqrt{\left(\frac{d_3}{d_0}\right)^2 + \alpha}$$

$$0 < \varepsilon \ll 1$$

die Ungleichungen (5.3-34) bis (5.3-36) erfüllen.

Da nach (5.3-37a)

$$\beta_{\min} < \beta < \beta_{\max} \quad \text{gilt,}$$

werden die Ungl. (5.3-35) und (5.3-36) stets erfüllt.

Nachzuweisen bleibt noch, daß die Ungl.

$$\left(d_0 + \frac{d_1}{2}\right) \left(\alpha d_0 + \frac{d_1}{2}\right) + d_3 (\beta d_0 + d_3) > 0 \quad (5.3-34)$$

erfüllt wird.

Da nach (5.3-27)

$$d_0, d_1 > 0, \quad d_3 \geq 0 \quad (A2-1)$$

ist, gilt

$$\left(d_0 + \frac{d_1}{2}\right) \left(\alpha d_0 + \frac{d_1}{2}\right) + d_3 (\beta d_0 + d_3) >$$

$$\left(d_0 + \frac{d_1}{2}\right) \left(\alpha d_0 + \frac{d_1}{2}\right) + d_3 (\beta_{\min} d_0 + d_3).$$

Falls nun

$$\left(d_0 + \frac{d_1}{2}\right) \left(\alpha d_0 + \frac{d_1}{2}\right) + d_3 (\beta_{\min} d_0 + d_3) \geq 0 \quad (A2-2)$$

ist, wird sichergestellt, daß (5.3-34) stets erfüllt wird.

Nach (5.3-37b) ist

$$\beta_{\min} = \frac{d_1 d_3}{d_o^2} - \frac{d_1}{d_o} \sqrt{\left(\frac{d_3}{d_o}\right)^2 + \alpha} \quad (\text{A2-3})$$

$$\text{Mit } y = \sqrt{\left(\frac{d_3}{d_o}\right)^2 + \alpha} = \sqrt{x^2 + \varepsilon} > 0 \quad (\text{A2-4})$$

$$\text{ergibt sich } \alpha = y^2 - \left(\frac{d_3}{d_o}\right)^2 \quad (\text{A2-5})$$

Mit den Abkürzungen

$$\left. \begin{aligned} b &= \frac{d_3}{d_o} \geq 0 \\ \text{und } a &= \frac{d_1}{2d_o} > 0 \end{aligned} \right\} \quad (\text{A2-6})$$

erhält man aus (A2-4)

$$y = \sqrt{b^2 + \alpha} > 0 \quad , \quad (\text{A2-7})$$

aus (A2-3)

$$\beta_{\min} = 2a(b - \sqrt{b^2 + \alpha}) = 2a(b - y) \quad (\text{A2-8})$$

und aus (A2-7)

$$\alpha = y^2 - b^2 \quad (\text{A2-9})$$

Aus (A2-2) ergibt sich dann

$$d_o^2 [(1+a)(\alpha+a) + b(b+\beta_{\min})] \geq 0 \quad .$$

Da nach (A2-1) $d_o > 0$ ist, muß

$$(1+a)(\alpha+a) + b(b+\beta_{\min}) \geq 0$$

sein.

(A2-8) und (A2-9) eingesetzt, ergibt

$$\begin{aligned} & (1+a)(y^2 - b^2 + a) + b[b + 2a(b-y)] = \\ & = (1+a)y^2 - 2aby + a + a^2 + ab^2 \geq 0. \end{aligned} \quad (\text{A2-10})$$

Wegen (A2-6): $a > 0$, $b \geq 0$ und

$$\begin{aligned} & 4(1+a)(a^2 + a + ab^2) = \\ & = 4[a^2 + a + ab^2 + a^3 + a^2 + a^2b^2] > 4a^2b^2, \end{aligned}$$

hat die Gleichung

$$f(y) = (1+a)y^2 - 2aby + a^2 + a + ab^2 = 0$$

keine reelle Lösung.

Da zugleich $f(0) > 0$

und $f(\infty) > 0$ gilt, folgt daraus, daß die Ungl. (A2-10) und damit auch (A2-2) stets erfüllt wird.

Das bedeutet, daß die Ungl. (5.3-34) bis (5.3-36) durch die Transformationsparameter α, β nach (5.3-40) und (5.3-41) erfüllt werden.

Literatur:

- | 1 | Schüßler, H.W.:
Digitale Systeme zur Signalverarbeitung
Springer-Verlag, Berlin 1973

- | 2 | Steffen, P.:
Ein Beitrag zur Theorie und Anwendung digitaler
Systeme mit mehreren Ein- und Ausgängen
Ausgewählte Arbeiten über Nachrichtensysteme Nr.31,
Hrsg. W. Schüßler, Erlangen

- | 3 | Mersereau, R.M.; Dudgeon, D.E.:
Two-Dimensional Digital Filtering
Proceedings of the IEEE, Vol.63, No.4, April 1975,
S.610-623

- | 4 | Schüßler, H.W.:
A 1-D Approach to 2-D Signal Processing
Symposium Bad Neuenahr, 1978
veröffentlicht in 'Advances in Digital Image Proces-
sing', Hrsg.: P. Stucki, IBM-Press, 1979

- | 5 | Huang, T.S.:
Picture Processing and Digital Filtering
Springer Verlag, Berlin 1975

- | 6 | Rabiner, L.R.; Gold, B.:
Theory and Application of Digital Signal Processing
Prentice-Hall, New Jersey, 1975

- | 7 | Rosenfeld, A.:
Digital Picture Analysis
Springer Verlag, Berlin, Heidelberg, New York 1976

- | 8 | Pratt, W.K.:
Digital Image Processing
John Wiley, New York 1978

- | 9 | Sandberg, I.W.:
Floating-point round-off accumulations in digital
filter realization
Bell Sys. Tech. J., Vol.46, S.1775-1791, Oct. 1967

- | 10 | Ni, M.D.; Aggarwal, J.K.:
Error analysis of two-dimensional recursive digital
filters employing floating-point arithmetic
IEEE Trans. Computers, Vol.C-25, No.7, S.755-764 1976

- | 11 | Maria, G.A.; Fahmy M.M.:
Bounds for the Amplitude of the Quantization Error
in Forced 1st-order 2-Dimensional Digital Filters
Circuit Theory and Applications, Vol.5, S.227-233, 1977

- | 12 | El-Agizi, N.G.; Fahmy M.M.:
Sufficient Conditions for the Nonexistence of Limit
Cycles in Two-Dimensional Digital Filters
IEEE Transactions on Circuits and Systems
Vol.CAS-26, No.6, S.402-406, 1979

- | 13 | Mc Clellan, J.H.:
The Design of Two-Dimensional Digital Filters by
Transformations
Proc. 7th Annual Princeton Conference on Information
Sciences and Systems, S.247-251, 1973

- | 14 | Mecklenbräucker, W.F.G., Mersereau, R.M.:
Efficient Design and Implementation for Two-Dimensional
FIR Digital Filters Designed using generalized McClellan
Transformation.
EASCON, S. 185A-185F, 1975

- | 15 | Rabiner, L.R.; Mc Clellan, J.H.; Parks, T.W.:
FIR Digital Filter Design Techniques Using Weighted
Chebyshev Approximation
Proc. of the IEEE, Vol.63, pp.595-610, Apr. 1975

- | 16 | Mersereau, R.M.; Mecklenbräucker, W.F.G.; Quateri, T.F.:
Mc Clellan transformations for two-dimensional digital
filtering: Design
IEEE Trans. Circuits Syst., S.405-414, July 1976

- | 17 | Razzak, M.A.; Cuthbert L.G.:
Two-dimensional nonrecursive digital filters designed
to arbitrary phase specifications
Electronic Circuits and Systems, Vol.3, No.5,
S.219-224, Sept.1979

- | 18 | Reed, I.S.; Troung, T.K.; Kwoh, Y.S.; Hall, E.L.:
Image processing by transforms over a finite field
Trans. on Computers, Vol.C-26, No.9, 1977

- | 19 | Rader, C.M.:
On the Application of the Number Theoretic Methods
of High-Speed Convolution to Two-Dimensional Filtering
Electronic Letters, S.575, 1974

- | 20 | Digital Signal Processing Comittee (Ed.),
Programs for Digital Signal Processing, IEEE Press,
New York, 1979

- | 21 | Eklundh, J.O.:
Fast Computer Method for Matrix Transposing
Transactions on Computers, S.801-803, 1972

- | 22 | Twogood, R.E.; Ekstrom, M.P.:
An Extension of Eklundh's Matrix Transposition and
its Application in Digital Image Processing
IEEE Transactions on Computers, Sept.1976

- | 23 | Eklundh, J.O.:
Efficient Matrix Transposition with Limited High
Speed Storage
FOA-Reports, Vol.12, No.1, 1978

- | 24 | Brandstätter, K.:
Programm für die zweidimensionale FFT zur Bildverar-
beitung unter Verwendung des Plattenspeichers an der
PDP 11/45
Studienarbeit am Lehrstuhl für Nachrichtentechnik
Erlangen, 1979

- [25] Harris, D.B.; Mc Clellan, J.H.; Chan D.S.K; Schüßler, H.W.:
Vector Radix Fast Fourier Transform
IEEE International Conference on ASSP, 1977
- [26] Mc Clellan, J.H.:
Recent Trends in Digital Signal Processing
Internat. Seminar, Zürich, Januar 17-21, 1977
- [27] Costa, J.M.; Venetsanopoulos, A.N.:
A Group of Linear Transformations for Two-Dimensional
Digital Filters
IEEE Transactions on Acoustics, Speech and Signal
Processing, S.424-425, Oct.1976
- [28] Shanks, J.L.; Treitel, S.; Justice, J.H.:
Stability and Synthesis of Two-Dimensional Recursive
Filters
IEEE Trans. on Audio- and Electroacoustics
Vol. AU-20, No.2, 1972
- [29] Karivaratharajan, P.; Swamy, M.N.S.:
Quadrantal Symmetry Associated with Two-Dimensional
Digital Transfer Functions
IEEE Transactions on Circuits and Systems, Vol.CAS-25,
No.6, S.340-343, June 1978
- [30] Goodmann, D.:
Some Stability Properties of Two-Dimensional Linear
Shift-Invariant Digital Filters
IEEE Trans. on Circuits and Systems, Vol.CAS-24,
No.4, April 1977
- [31] Shaw, G.A.; Mersereau, R.M.:
Design, Stability and Performance of Two-Dimensional
Recursive Digital Filters
Technical Report E21-BO5-1, Georgia Institute of
Technology, School of Electrical Engineering, Atlanta,
Georgia, Dec.1979

- [32] Goodman, D.:
Some Difficulties with the Double Bilinear Transformation in 2-D Recursive Filter Design
Proceedings of the IEEE, Vol.66, No.7, S.796-797, 1978
- [33] Shaw, G.A.; Mersereau, R.M.:
Design of Two-Dimensional Recursive Filters with Arbitrary Support using Quarter-Plane Design Algorithms
Proceedings of IEEE, Vol.67, No.7, S.1081-1083, July 1979
- [34] Huang, T.S.:
Stability of Two-Dimensional Recursive Filters
IEEE Trans. Audio Electroacoustics, Vol.AU-20, S.158-163, 1972
- [35] Murray, J.:
Another Proof and a Sharpening of Huang's Theorem
IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-25, 1977
- [36] Davis, D.L.:
A Correct Proof of Huang's Theorem on Stability
IEEE Trans. Acoust., Speech, Signal Processing, S.425-427, Oct.1976
- [37] Goodman, D.:
An Alternate Proof of Huang's Stability Theorem
IEEE Trans. Acoust., Speech, Signal Processing, S.426-427, Oct.1976
- [38] De Carlo, R.A.; Murray, J.; Saeks, R.:
Multivariable Nyquist Theory
Int. J. Control, 1977, Vol.25, No.5, S.657-675
- [39] Strintzis, M.G.:
Test of Stability of multidimensional Filters
IEEE Trans. Circuits Syst., Vol.CAS-24, S.432-437 Aug. 1977

- | 40 | Ekstrom, M.P.:
Design of 2-D Digital Filters
Seminar am Lehrstuhl für Nachrichtentechnik, Erlangen,
WS 78/79
- | 41 | O'Connor, B.T.; Huang, T.S.:
Stability of general twodimensional recursive digital
filters
IEEE Trans. Acoust., Speech, Sign. Proc., Vol.ASSP-26,
S.550-560, Dec.1978
- | 42 | Ekstrom, M.P.; Woods, J.W.:
Two-Dimensional Spectral Factorization with Applications
in Recursive Digital Filtering
IEEE Trans. Acoust., Speech, Sign. Proc., Vol.ASSP-24,
No.2, S.115-118, April 1976
- | 43 | Maria G.A.; Fahmy M.M.:
On the Stability of Two-Dimensional Digital Filters
IEEE Trans. Audio and Electroacoustics, S.470-472,
Oct.1973
- | 44 | Anderson, B.D.; Jury, E.I.:
Stability Test for Two-Dimensional Recursive Filters
IEEE Trans. Audio and Electroacoustics, Vol.AU-21,
No.4, S.366-372, Aug.1973
- | 45 | Siljak, D.D.:
Stability criteria for two-variable polynomials
IEEE Trans. Circuits Syst., Vol.CAS-22, S.185-189,
March 1975
- | 46 | Ekstrom, M.P.; Twogood, R.E.:
Design of stable 2-D Half-Plane Recursive Filters
using Spectral Factorization
European Conference on Circuit Theory and Design,
Lausanne, S.351-355, 1978

- [47] Ekstrom, M.P.; Twogood, R.E.; Woods, J.W.:
Design of Stable 2-D Half-Plane Recursive Filters
using Spectral Factorization
International Conference on Acoustics, Speech and
Signal Processing, Tulsa Oklahoma, April 10-12, 1978

- [48] Ekstrom, M.P.; Twogood, R.E.:
A Stability Test for 2-D Recursive Digital Filters
using the Complex Cepstrum
Proc. 1977 Int. Conf. Acoust., Speech, Sign. Proc.,
S.535-538, April 1977

- [49] Shaw, G.A.:
An Algorithm for Testing Stability of Two-Dimensional
Digital Recursive Filters
Proc. of Int. Conf. ASSP, S.769-772, 1978

- [50] Goodmann, D.M.:
A Design Technique for Circularly Symmetric Low-
Pass Filters
IEEE Trans. on Acoustics, Speech and Signal Proc.,
Vol.ASSP-26, No.4, S.290-304, Aug.1978

- [51] Steffen, P.:
Stabilität eines allg. 2-D Blocks 2.Grades
Persönliche Mitteilungen

- [52] Dudgeon, D.E.:
The existence of cepstra for two-dimensional rational
polynomials
IEEE Trans. Acoust., Speech and Signal Processing,
Vol.ASSP-23, S.242-243, April 1975

- [53] Pistor, P.:
Stability Criterion for Recursive Filters
IBM, J. Res. Develop., S.59-71, Jan.1974

- [54] Murray, J.J.:
Spectral Factorization and Quarter-Plane Digital
Filters
IEEE Transactions on Circuits and Systems, Vol.CAS-25,
No.8, S.586-592

- [55] Tribolet, J.M.:
A new phase unwrapping algorithm
IEEE Trans. Acoust. Speech and Sign. Proc., Vol.ASSP-25,
S.170-177, April 1977

- [56] Jury, E.I.:
Theory and Application of the Z-Transform Method
John Wiley, New York 1967

- [57] Entenmann, W.:
Optimierungsverfahren
Hüthig-Verlag, Heidelberg, 1976

- [58] Möhringer, P.:
Optimierungsprogramm mit dem Simplex-Verfahren
Persönliche Mitteilungen

- [59] Czarnach, R.:
Realisierung von 1-D rekursiven Nullphasenfiltern
Persönliche Mitteilungen

- [60] Ekstrom, M.P.; Twogood, R.E.:
Why Filter recursively in two dimensions?
Proc. 1979 IEEE Conf. Acoust. Speech, Sign. Proc.,
S.20-23, April 1979

- [61] Betran, M.S.:
Approximation of digital filters in one and two
dimensions
IEEE Trans. Acoust. Speech, Sign. Proc., Vol.ASSP-23
S.438-443, Oct.1975

- [62] Nowrouzian, B.; Ahmadi, M.; King, R.A.:
On the space-domain design techniques of n-dimensional recursive digital filters
Proc. 1977 IEEE Int. Conf. Acoust. Speech, Sign. Proc.,
S.531-534, April 1977
- [63] Parker, S.R.; Souchon, L.:
Synthesis of N-Dimensional Recursive Digital Filters
by Taylor Series Expansion
IEEE Circuits and Systems, Vol.24(1), S.28-34, 1977
- [64] Bordner, G.W.:
Time-domain design of stable recursive digital filters
Ph. D. Thesis, State University of New York, Electrical
Engineering Dept., Buffalo, New York, 1974
- [65] Maria, G.A.; Fahmy, M.M.:
An lp Design Technique for Two-Dimensional Digital
Recursive Filters
IEEE Trans. on Acoust., Speech and Sign. Proc.,
S.15-21, 1974
- [66] Pendergrass, N.A.:
Comments on "An lp Design Technique for Two-Dimensional
Digital Recursive Filters"
IEEE Trans. on Acoust., Speech and Sign. Proc.,
S.389-390, 1975
- [67] Bernabo, M.; Cappellini, V.; Emiliani, L.:
A Method for designing 2-D Recursive Digital Filters
having a circular symmetry
Electronic Letters, 12(11), S.288-289, 1976
- [68] Ramamoorthy, P.A.; Bruton, L.T.:
Design of Stable 2-Dimensional Discrete Recursive
Filters
Electronic Letters, Vol.12, No.25, S.659-660, 1976
Comments and Corrections
Electronic Letters, Vol.13, No.16, S.491-492, 1977

- |69| Twogood, R.E.; Mitra, S.K.:
Computer-Aided Design of Separable Two-Dimensional
Digital Filters
Vol.ASSP-25, No.2, S.165-169, 1977

- |70| Chakrabarti, S.; Mitra, S.K.:
Design of 2-Dimensional Digital-Filters via
Spectral Transformations
Proceedings of IEEE, 65(6), S.905-914, 1977

- |71| Prasad, K.P.; Reddy, P.S.:
Method for Designing Stable 2-D Digital Filters
Electronic Letters, Vol.13, No.13, S.385-387, 1977

- |72| Costa, J.M.; Venetsanopoulos, A.N.:
Design of Circularly Symmetric Two-Dimensional
Recursive Filters
IEEE, ASSP-22, No.6, S.432-443, 1974

- |73| Pendergrass, N.A.; Mitra, S.K.; Jury, E.I.:
Spectral Transformations for Two-Dimensional Digital
Filters
IEEE Circuits and Systems, Vol.CAS-23, No.1, S.26-35
1976

- |74| Chang, H.; Aggarwal, J.K.:
Design of Two-Dimensional Recursive Filters by
Interpolation
IEEE Trans. on Circuits and Systems, Vol.CAS-24,
No.6, S.281-291, 1977

- |75| Aly, S.A.H.; Fahmy, M.M.:
Design of two-dimensional digital filters with
specified magnitude and group delay characteristics
IEEE Trans. Circuits and Systems, Vol.CAS-25,
S.908-915, 1978

- [76] Delcaro, L.G.; Sicuranza, G.L.:
Design of two-dimensional recursive digital filters
IEEE Trans. Acoust., Speech, Sign. Proc., Vol.ASSP-25,
S.577-579, Dec.1977
- [77] Ekstrom, M.P.; Twogood, R.E.; Woods, J.W.:
Two-Dimensional Recursive Filter Design - A Spectral
Factorization Approach
IEEE Trans. Acoust., Speech, Sign. Proc., Vol.ASSP-28,
No.1, Febr.1980
- [78] Garriboto, G.:
A new approach to half-plane recursive filter design
ASSP-Workshop on Two-Dimensional Digital Signal Pro-
cessing, Berkley, California, 3.-4. Oct. 1979
- [79] Chang, H.; Aggarwal, J.K.:
Design of two-dimensional semicausal recursive filters
IEEE Trans. Circ. and Systems, CAS-25, S.1051-1059, 1978
- [80] Deczky, A.G.:
Synthesis of recursive digital filters using the
minimum p-error criterion
IEEE Trans. Audio Electroacoust., Vol.AU-20,
S.257-263, Oct.1972
- [81] Jury, E.I.:
Inners and Stability of Dynamic Systems
John Wiley, New York, 1974
- [82] Dehner, G.F.:
Program for the Design of Recursive Digital Filters
in: Programs for Digital Signal Processing, Kap.6.1,
IEEE-Press, New York, 1979

- | 83 | Ahmadi, M.; Constantinides, A.G.; King, R.A.:
Design Technique for a class of stable Two-Dimensional Recursive Digital Filters
IEEE Internat. Conf. on Acoust., Speech and Sign. Proc., April 12-14, S.145-147, Philadelphia, 1976
- | 84 | Ali, A.M.:
Design of Inherently Stable Two-Dimensional Recursive Filters Imitating the Behaviour of One-Dimensional Analog Filters
Proc. of Int. Conf. ASSP, S.765-768, 1978
- | 85 | First ASSP-Workshop on Two-Dimensional Digital Signal Processing
Session D2, Chairman Mc Clellan, J.H.:
Comparison of Recursive and Nonrecursive Filters
Berkley, California, Oct.3-4, 1979
- | 86 | Bernstein, R.:
Ein Beitrag zum Entwurf von 2-D-IIR-Filtern mit zirkularsymmetrischem Betragsfrequenzgang
DFG-Kolloquium, Göttingen, Febr.1980
- | 87 | Modestino, J.W.; Fries, R.W.:
Edge Detection in Noisy Images using Recursive Digital Filtering
Computer Graphics and Image Processing 6, S.409-433, 1977
- | 88 | Treitel, S.; Shanks, J.L.; Frasier, C.W.:
Some Aspects of Fan Filtering
Geophysics, Vol.32, S.789-800, Oct.1967
- | 89 | Fettweis, A.:
Zweidimensionale Wellendigitalfilter mit fächerförmiger Übertragungsfunktion
DFG-Kolloquium, Göttingen, Febr.1980

- [90] Fettweis, A.:
Design of Recursive Quadrant Fan Filters
AEÜ, Band 34, S. 97-103, März 1980
- [91] Fail, J.P.; Grau, G.:
Les Filtres en éventail
Geophysical Prospecting, Vol.XI, S.131-163, 1963
- [92] Mc Clellan, J.H.; Parks, T.W.:
Equiripple approximation of fan filters
Geophysics, Vol.37, S.573-583, Aug. 1972
- [93] Kulhanek, O.:
Introduction to Digital Filtering in Geophysics
Elsevier Scientific Publishing Co., Amsterdam 1976
- [94] Zimmer, H.G.; Kronberg, H.; Bernstein, R.;
Neuhoff, V.:
Improvements in microphotometry by digital signal-
processing
Pattern Recognition, Vol.13, S.79-82, Jan. 1981
- [95] Zimmer, H.G.; Neuhoff, V.:
Quantitative Auswertung von zweidimensionalen Mikro-
Chromatogrammen
Informatik Fachberichte 8(Hrsg. H.H.Nagel), Digital
Image Processing 12, Springer-Verlag, Berlin, Heidel-
berg, New York 1977
- [96] Zimmer, H.G.; Kronberg, H.; Neuhoff, V.:
Quantitative evaluation of Chromatograms
Proceedings of the Fourth International Joint Conference
on Pattern Recognition and Image Processing Publication
78 CH 1318-5C, S.834-836, Kyoto 1978
- [97] Kronberg, H.; Zimmer, H.G.; Neuhoff, V.:
Segmentierung in digitalen Chromatogrammbildern mit
abgestufter Rastergröße
Informatik-Fachberichte 20, Angewandte Szenenanalyse
(Hrsg. J.P. Foith), S.254-259

|98| Frank, G.:

Untersuchung von linearen und nichtlinearen Operatoren
zur Kantendetektion in Bildern

Diplomarbeit am Lehrstuhl für Nachrichtentechnik

Erlangen, 1980

L e b e n s l a u f

29.5.1946	geboren in Herrieden/Mittelfranken
1952 - 1959	Volksschule, Ansbach
1959 - 1963	Gewerberealschule, Nürnberg
1963	Mittlere Reife
1963 - 1965	Praktikum bei der Firma Neumeyer (Kabelmetall) und der Deutschen Bundespost in Nürnberg
1965 - 1968	Studium der Elektrotechnik am Ohm-Polytechnikum in Nürnberg
1968	Studienabschluß als Ing.grad., fachgebundene Hochschulreife
1968	Laboringenieur bei der Firma Hydrometer in Ansbach
1968 - 1969	Studium der Elektrotechnik an der Universität Erlangen-Nürnberg (1. Semester)
1969 - 1970	Grundwehrdienst als Wehrpflichtiger
1970 - 1974	Fortsetzung des Elektrotechnik-Studiums an der Universität Erlangen-Nürnberg
1974	Abschluß des Studiums mit dem Diplom
1974 - 1975	Tätigkeit als Entwicklungsingenieur bei der Firma Siemens (UB Medizinische Technik) in Erlangen
seit 1975	Wissenschaftlicher Assistent am Institut für Nach- richtentechnik der Universität Erlangen-Nürnberg

Copy-Center 2000

852 ERLANGEN · Obere Karlstr. 30
Tel. 09131-26197